Single-cell approaches reveal functional and molecular heterogeneity in malignant haematopoietic stem cells

Mairi Stella Shepherd
Homerton College

UNIVERSITY OF CAMBRIDGE

Department of Haematology
University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

August 2020

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the School of Clinical Medicine Degree Committee.

Single-cell approaches reveal functional and molecular heterogeneity in malignant haematopoietic stem cells

Mairi Stella Shepherd

Abstract

Recent advances in single-cell technologies have permitted the investigation of heterogeneous cell populations at previously unattainable resolution. In this thesis single-cell approaches are applied to resolve the molecular mechanisms driving disease in mouse haematopoietic stem cells (HSCs), using JAK2 V617F mutant myeloproliferative neoplasms (MPNs) as a model. This study utilises single-cell gene expression and functional assays to identify a subset of JAK2 V617F mutant HSCs that display defective self-renewal. This defect is rescued by crossing JAK2 V617F mice with mice lacking TET2, the most commonly co-mutated gene in patients with MPN. Single-cell gene expression profiling of JAK2 V617F-mutant HSCs revealed reduced expression of specific self-renewal regulator genes, some of which were restored to normal levels in single TET2/JAK2 mutant HSCs. Of these, *Bmi1* and, to a lesser extent, *Pbx1* and *Meis1* overexpression in JAK2-mutant HSCs could improve stem cell self-renewal, allowing development of a disease phenotype in functional assays. Together, these findings refine the molecules involved in clonal expansion of MPNs, and highlight the power of single cell approaches in deconstructing the functional characteristics and molecular network of normal and malignant stem cells.

Preface

Chapter 3

All work was carried out by me, under the supervision of David Kent, with the exception of the human data in section 3.3 which was generated by Christina Ortmann and David Kent. All single cell sorts were carried out by the Flow Cytometry Core Facility at the Cambridge Institute for Medical Research (CIMR). The CBP, NMP1, CALR and p53 mouse sort preparation was performed with the assistance of Nicola Wilson

Chapter 4

All work was carried out by me, under the supervision of David Kent, with the exception of the additional transplantation data that was added for statistical comparison in section 4.2.2, which was from an experiment performed by Juan Li.
Mouse transplantation injections were performed by David Kent, Dean Pask, and Caroline Oedekoven, serial bleeds were performed by me after receiving training from Tina Hamilton and Dean Pask.

Chapter 5

All work was carried out by me, under the supervision of David Kent, with the exception of the multiplexed qPCR which was performed by Jiangbing Li and David Kent, analysis of this data was performed by Nicola Wilson. Mouse transplantation injections were performed by David Kent, and Caroline Oedekoven. Mouse cell sorts were carried out by the Flow Cytometry Core Facility at CIMR. Human cell sorts were carried out by Cancer Research UK Cambridge Institute flow cytometry facility.

Acknowledgments

It would not have been possible to write this thesis without the help and support of many people. I would like to express my gratitude to my supervisor David Kent for his guidance, support, and advice throughout this PhD. I am also very grateful for the kindness, support, and encouragement of my other supervisors Brian Huntly and Armin Sepp. This research would not have been possible without my funding bodies, BBSRC and GSK, thank you to them for sponsoring my research and tutelage.

I would like to acknowledge the invaluable input of Juan Li, for her wide-ranging guidance and generosity with her expertise, and Nicola Wilson for her advice and for analysing the multiplexed qPCR data. Thanks are also owed to Priyanka Tibarewal and Joaquina Delas Vivas for their assistance in lentivirus production, and Jiangbing Li for carrying out the multiplexed qPCR on the TET2 knock-out and double mutant samples. I would also like to acknowledge Reiner Schulte, Chiara Cossetti, Gabriela Grondys-Kortaba, and Michal Maj in the Flow Cytometry Facility of the Cambridge Institute for Medical Research, and Richard Grenfell in the Cancer Research UK Cambridge Institute sort facility for technical assistance and suggestions. For their assistance with mouse work I would like to thank Dean Pask, Tina Hamilton and Caroline Oedekoven, and the Central Biomedical Services staff for animal housing and care.

These acknowledgments would not be complete without mentioning my lab pals, especially Nina Øbro, Miriam Belmonte, Rebecca Caeser, Craig McDonald and James Che. It was a great pleasure working with them, and I appreciate their help, ideas and humour. Last but not the least, I would like to thank my parents, grandparents, my brother and sister, and Nick Smith for supporting me throughout writing this thesis and throughout life in general.

Abbreviations

AGM; Aorta-gonads-mesonephros
ALL; Acute lymphoblastic leukaemia
AML; Acute myeloid leukaemia
APC; Allophycocyanin
ARCH; Age-related clonal haematopoiesis
ASXL; Additional sex combs like
BFUe; Burst-Forming Unit erythrocytes
BM; Bone marrow
BV; Brilliant Violet
CAFC; Cobblestone area–forming cell
CALR; Calreticulin
CBP; CREB-binding protein
CD; Cluster of differentiation
cDNA; Complementary deoxyribonucleic acid
CFC; Colony forming cell
CFU-S; Colony-forming units – spleen
CLL; Chronic lymphocytic leukaemia
CML; Chronic myelogenous leukaemia
CMML; Chronic myelomonocytic leukaemia
DNMT; DNA methyltransferase
ECM; Extracellular matrix
EPCR; Endothelial protein C receptor
EPO; Erythropoietin
ET; Essential thrombocythaemia
EV; Empty vector
EZH; Enhancer of zeste
FBS; Foetal bovine serum
FITC; Fluorescein isothiocyanate
FL; Foetal liver
G-CSF; Granulocyte colony-stimulating factor
GM; Granulocyte/macrophage
HCT; Haematocrit
HDAC; Histone deacetylase
HGB; Haemoglobin
HSC; Haematopoietic stem cell
HSPC; Haematopoietic stem and progenitor cell
IDH; Isocitrate dehydrogenase
IFC; Integrated fluidics chip
IFN; Interferon
IL; Interleukin
JAK; Janus kinase
Lin; Lineage markers
LSK; Lineage$^-$ Sca1$^+$ c-Kit$^+$
LTC-IC; Long-term culture-initiating cell
M-CSF; Macrophage colony-stimulating factor
MDS; Myelodysplastic syndromes
MF; Myelofibrosis
MNC; Mononuclear cell
MPN; Myeloproliferative neoplasms
MPP; Multipotent progenitor
mRNA; messenger ribonucleic acid
NM; Non-mutant
NPM; Nucleophosmin
PacB; Pacific Blue
PB; Peripheral blood
PBS; phosphate-buffered saline
PCA; Principal component analysis
PE; Phycoerythrin
pIpC; Polyinosinic-polycytidylic acid
pMF; Primary myelofibrosis
PRC; Polycomb repressive complex
PV; Polycythaemia vera
qPCR; Quantitative polymerase chain reaction
RBC; Red blood cell
RNAseq; Ribonucleic acid sequencing
SCF; Stem cell factor
scRNAseq; Single cell RNA sequencing
SEM; Standard error of the mean
sMF; Secondary myelofibrosis
STAT; Signal transducers and activators of transcription
T-ALL; T-cell acute lymphoblastic leukaemia
t-SNE; T-distributed stochastic neighbour embedding
TET; Ten-eleven translocation
TF; Transcription factor
TGF; transforming growth factor
THPO; Thrombopoietin
TKI; Tyrosine kinase inhibitor
TNF; tumour necrosis factor
WT; Wild type
αKG; Alpha-ketoglutarate
2HG; 2-Hydroxyglutarate
5-hmCyt; 5-Hydroxymethyl-cytosine
5-mCyt; 5-Methyl-cytosine
7AAD; 7-Aminoactinomycin D

Table of Contents

Table of Figures

Table of Tables

1

Introduction

1.1 Haematopoiesis and HSCs

Blood is a complex and multifaceted tissue: circulating through the body, it transports oxygen, nutrients and hormones; removes waste; regulates body temperature and pH; and provides a crucial defence from pathogens, viruses and injury. The blood is made up of an aqueous component (plasma), and a cellular component; the cellular component of blood contains a diverse array of cell types which reflect the blood's many functions. For example, erythrocytes carry oxygen to, and carbon dioxide from, tissues. Platelets play a key role in blood clotting in response to vessel damage. B and T lymphoid cells are the primary players in the adaptive immune system responsible for acquired defence against pathogens. While the innate immune system – encompassing lymphoid cells such as natural killer and mast cells, and phagocytic cells such as macrophages and neutrophils – is involved in non-specific pathogen defence.[1]

A sufficient number and balance of each blood cell type needs to be maintained to efficiently fulfil the blood's numerous functions. However, as the mature cells of the blood have relatively short life spans,[1] and cannot self-renew, they need to be continually replaced by more potent progenitor cells. The process of blood cell production is termed haematopoiesis, and it begins with the haematopoietic stem cell (HSC). HSCs are a rare cell population that sit at the top of the haematopoietic hierarchy, the most potent cell in the blood. HSCs are defined by two fundamental characteristics: the ability to differentiate into multiple mature blood cell lineages (via a series of progenitors and substantial amplification in cell number), and the ability to

self-renew to create equally potent daughter HSCs, to sustain the population of immature multipotent stem cells. As a population, HSCs must undergo a number of fate decisions on a cell-by-cell basis, such as whether or not to self-renew or differentiate, divide by symmetric or asymmetric division, or to undergo cell death. All of these fate decisions are coordinated by molecular changes and are balanced within the population of HSCs to ensure equilibrium between all mature blood lineages and a sufficient population of HSCs to sustain haematopoiesis for the lifetime of the organism.[2]

Understanding the haematopoietic system has been a longstanding goal of the scientific and medical community for several reasons. Firstly, the use of blood and blood fractions, like platelets, plasma or red blood cells, are in common use in the clinical setting, for example for blood transfusions in surgery or after childbirth, or treatment of anaemias, cancer and blood disorders. The current source of blood and blood fractions is from healthy volunteer donors and is often insufficient for the demand. HSC transplantation is the only curative option for many diseases, but the availability of donor HSCs is limited. The ability to generate these products *in vitro* would allow for a larger and more homogenous source of blood cells, which could be generated to fit demand and tailored to the required cell and histocompatibility types. Likewise, by gaining a deeper understanding of blood diseases such as leukaemia, new therapies could be generated, either for diseases that respond poorly to current therapies, or to replace therapies that are poorly tolerated in patients. Finally, by understanding how the blood carries out its functions, like immune control, blood cells could be manipulated to allow new and highly specific therapies, such as chimeric antigen receptor T (CAR-T) cells.[3] By understanding HSCs more fully, and especially the process by which they self-renew, options such as expansion for transplantation and gene therapy could become a real clinical option.

1.2 Identification of HSCs

The HSC was first alluded to experimentally during studies on the effects of ionizing radiation. Jacobsen et al. (1949) found that shielding the spleen (or femur) of mice with lead protected them from the lethality of irradiation.[4] This led Lorenz et al. (1951) to discover that transplantation of bone marrow (BM) post-irradiation could protect against irradiation lethality.[5] It was revealed that this protection was due to cells (rather

than injection stimulation), by Barnes and Loutit (1954)[6], and Main and Prehn (1955),[7] thereby setting the stage for the study of the cells of the BM.

Till and McCulloch made the first functional description of the haematopoietic stem cell – the first adult stem cell functionally demonstrated to exist – in 1961.[8] They discovered that when BM is transplanted from one mouse to an irradiated recipient 'nodules' of rapidly dividing haematopoietic tissue appeared on the spleen. Mature cells could be identified within these nodules and some colonies contained multiple lineages. The authors found that the number of nodules that appeared on the spleen was directly proportional to the number of cells transplanted and referred to these cells with the capacity to produce colonies as colony-forming units – spleen (CFU-S).[8] These colonies were proven to be of single cell origin by inducing mutations in the initiating cells via low level irradiation and studying the karyotype of the resulting colonies.[9] The full capacity of a single HSC - completely on its own - to repopulate a mouse was recognised in 1996 when the first single cell transplantations of functional HSCs was published,[10] formally demonstrating that a single cell ($CD34^{-/low}$, c-Kit^{+}, $Sca1^{+}$, Lin^{-}) could successfully repopulate the BM of 21% of recipients, with reconstitution in the lymphoid and myeloid compartment in 85% of successful transplants.

Following the early HSC experiments of the 1960s, focus moved quickly to isolating these cells, to permit the unique and defining characteristics of HSCs to be studied. Isolation strategies ranged across research groups, with some opting to isolate HSCs based on their functional properties, for example Rhodamine 123[11] or Hoescht 33342 dye exclusion,[12] resistance to 5-fluorouracil,[13] or resistance to gamma-irradiation.[14] Others sought to isolate stem cells based upon various antibodies binding to cell surface proteins. Neither antibody binding nor dye exclusion methods for HSC isolation would have been possible without the popularisation of fluorescence activated cell sorting. The first HSC sorting strategy that utilised cell surface marker properties was published by Visser et al. (1984).[15] This group first performed density gradient centrifugation and sorted the low-density cells for WGA (wheat germ agglutinin) binding positivity, then resorted for high expression of H-2K (histocompatibility antigen). When assessed in a CFU-S assay the group reported an 135 fold enrichment of HSCs over non-purified BM.

In 1988 Spangrude, Heimfield and Weissman[16] improved upon the HSC sorting strategy. By isolating $Lin^{-}Thy1^{lo}Sca1^{+}$ cells the group were able to increase the efficiency of spleen or thymic colony formation by 1000-fold over non-purified BM.[16] Subsequently a number of additional HSC and progenitor markers have been identified the most notable of which being c-Kit,[17] Endothelial protein C receptor (EPCR),[18] cluster of differentiation (CD)49b,[19] the SLAM markers[20] and CD34/Flk2.[21] A number of different flow cytometric sorting strategies have been developed allowing the isolation of populations from fresh BM of young adult mice with HSC frequencies of >50%.[20,22–25] Currently techniques like index sorting are being used to refine these sorting strategies further to obtain even higher HSC purities; flow cytometric index sorting records the fluorescence intensity of each immunophenotypic marker, thereby allowing the retrospective coupling of markers and functional output. This has been combined with single cell functional *in vitro*[26] and *in vivo*[27] assays to refine the effectiveness of sort panels.

Index sorting is also used in the refinement of human HSCs sorting panels,[28] which remain less well optimised that murine HSC isolation. Currently the best human HSC sorting strategy isolates HSCs at a frequency of ~1 in 10 from cord blood (with lower frequencies in BM and peripheral blood).[29] This, in combination with the variability between human individuals' genetic, clinical and lifestyle factors, and mutation status varying between cells, mean that mouse models continue to be heavily relied upon for the study of HSC biology.

1.3 The changing properties of HSCs

1.3.1 HSCs throughout life

In mice, HSCs can first be detected at embryonic day (E)10.5 in the aorta-gonads-mesonephros (AGM) and aorta associated vessels (vitelline vessels and umbilical vessels).[30] HSCs form in the AGM from a specialised endothelium on the ventral side of the dorsal aorta called the haemogenic endothelium, in a process termed endothelial-to-haematopoietic transition.[30] The AGM is the site of *de novo* HSC production in the developing embryo, but it is not the site of HSC expansion or differentiation. The primary site of HSC expansion and differentiation in the developing foetus is the foetal

liver (FL). As soon as HSCs emerge, they colonise this site, before migrating to the BM prior to birth.[31]

The properties of HSCs in the FL differ from those of the adult BM in their proliferation rate, self-renewal capacity and lineage output. Foetal HSCs are in active cell cycle and have a substantially increased capacity for self-renewal compared to adult HSCs, a trait that persists until 3-4 weeks after birth when a more quiescent adult phenotype is established.[32] The foetal HSC properties of increased proliferation rate and increased self-renewal capacity are the same properties that must be activated for clonal expansion in the cancer setting, it is therefore understandable that some of the molecular regulators involved in developing HSCs have also been implicated in cancer, some of these will be discussed later (in section 1.4.2).

During adult life it is important for the HSC population to maintain a balance between differentiation and self-renewal. Excessive differentiation or defective self-renewal leads to depletion of the HSC pool, whereas unrestrained self-renewal or lack of differentiation can result in cancer. Control of this delicate balance is regulated by a number of intrinsic factors (e.g. transcription factors and epigenetic regulators) and extrinsic factors (e.g. BM microenvironment: cytokines, cell:cell interactions etc).[33]

With age the balance between differentiation and self-renewal becomes is disrupted, leading to a decline in regenerative capacity.[34] In mouse, two major changes occur with age in the haematopoietic system: depletion of the stem cell pool, and lineage skewing of HSC progeny. As mice age the number of phenotypic HSCs increases,[35–37] however the long-term functional capacity of these cells, as measured by serial transplantation, decreases.[35,36] With regard to lineage output, in young mice a strict balance of cell types is maintained, however aged mouse HSCs tend to have a myeloid bias due to a reduced capacity to give rise to lymphoid cells.[36,38]

Numerous factors have been suggested as the cause of stem cell aging, such as genetic mutations, epigenetic changes, and changes in the microenvironment and stem cell niche.[37] The majority of the aged phenotype has been linked to intrinsic cell changes,[39] as demonstrated when aged HSCs are transplanted into a young microenvironment

resulting in expansion of the HSC population and reduced lymphoid capacity in young recipients.[40]

With the reduction in HSC self-renewal that aging brings, and the subsequent exhaustion of clones, the diversity of clones within the stem cell compartment decreases. This reduction in polyclonality has been observed though viral barcoding of HSCs; 'tagging' HSCs with a genetic barcode which is passed on to all the cells progeny, allowing the clonal dynamics of the stem cell compartment to be observed. This approach has shown that if young cells are transplanted and followed over time, grafts begin as very polyclonal but progressively decrease in the number of clones contributing to blood production as the mice age.[41] Barcoded tracking also highlights the heterogeneity of the HSC population in terms of cellular output, time to contribution, length of clone life, and number of cells produced.

1.3.2 Clonal Haematopoiesis

Defining a clone is a difficult task; with "a population of cells of single cell origin" representing a common definition. However, this could in fact refer to every organism - emerging from the single cell that is the fertilised egg.[42] In the cancer setting it is somewhat easier to define the clone where all cells share the same genetic aberration that was acquired in a single cell. Similarly, a clone is easily defined in cellular barcoding studies where a single cell is infected with a unique barcode. In the normal haematopoietic setting, as HSCs give rise, though significant progenitor amplification, to large numbers of mature cells, HSC "clones" are often referred to as the progeny of an HSC. These clones are very hard to define due to the fact that HSCs are giving rise to more HSCs and clones form inside clones, the line between these clones is impossible to draw. Clonal haematopoiesis refers to the relative over-representation of one clone within the mature cell output.

Clonal haematopoiesis can occur as a result of neutral drift; random chance leading to the expansion or exhaustion of clones. A analogy can be drawn between the expansion or disappearance of haematopoietic clones and family surnames in Western cultures where children retain the male surname; if one assumes an even distribution of family names within a town, through chance, some families will have sons carrying on the family names, others will have daughters who give up their family name. Over time

some family names will expand while others will diminish or disappear altogether without any sort of "selective pressure". The same is true of haematopoietic clones; the balance between differentiation and self-renewal will cause some clones to expand over time while others decrease in size, this is not due to outcompeting of one clone over another but neutral drift. However positive selection, and not neutral drift, has been shown to be the major driving force in clonal haematopoiesis,[43] there are a number of ways in which a clone can make up a larger proportion of the haematopoietic compartment beyond the levels of neutral drift (Figure 1)[44];

1. *The clone expands.*

If a clone gains an advantage that results in the dramatic growth the clone, leaving the rest of the haematopoietic cells unaffected, the overall size of the population is increased, the clone making up a large proportion of the population. The clone has not compromised the fitness of other clones but established its proportion of the haematopoietic population by expanding the population size.

2. *The clone outcompetes the other HSCs*.

The population size of the haematopoietic population remains stable but an increasing proportion of it is made up of the clone as it outcompetes and displaces the other HSCs. It is important to remember that in this case the clone only needs to be more competitive than the other HSCs, in the case of aging HSC are less competitive and therefore a mutation that may be insufficient to outcompete young HSCs may be sufficient to outcompete aged HSCs.

3. *Population decrease.*

With aging HSCs become less competitive and the population size decreases. With the reduction in size of other clones, one clone can take up an increasing amount of the haematopoietic space. This clone may not have any particular advantage, but as a consequence of decreasing population size it artificially appears to grow.

The presence of clonal haematopoiesis in human individuals increases with age; in individuals under 50 only 1% have clonal haematopoiesis as measured by whole exome

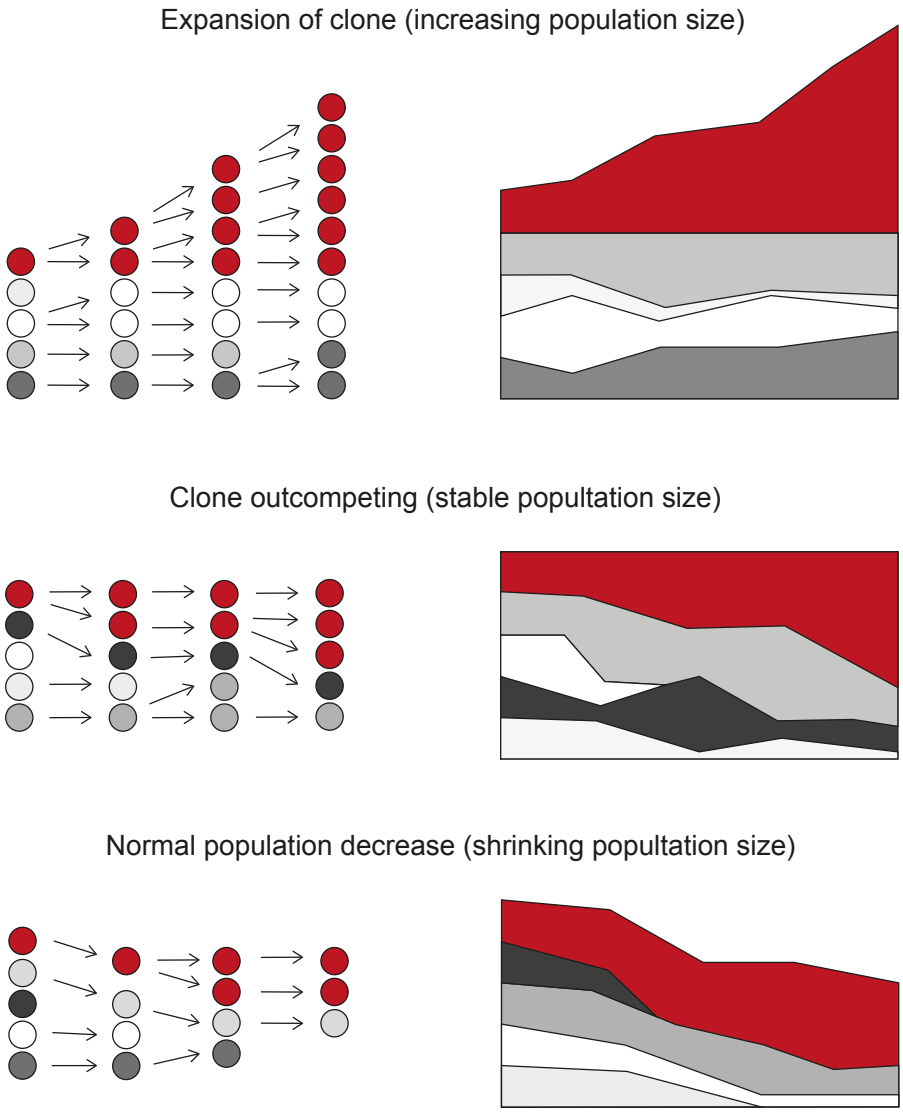


Figure 1 Means of acquiring clonal haematopoiesis.
As clonal haematopoiesis is measured as a proportion of the overall haematopoietic stem cell clone population there are a number of ways in which clonal haematopoiesis can occur; the clone can expand enlarging the population size, the clone can outcompete other clones keeping a stable population size, or the population itself can shrink.

sequencing, however, in individuals over 65 10% have observable clonal haematopoiesis.[45] Certain mutations, commonly found in myeloid malignancies, are associated with clonal haematopoiesis without clinical features, these mutations provide an advantage to the clone, allowing clonal expansion[46]. Age-related clonal haematopoiesis (ARCH) is defined as the expansion of HSPC clones, harbouring specific disruptive genetic variants, in individuals without a diagnosis of haematological malignancies.[47] The most commonly mutated genes in ARCH are *DNMT3A, ASXL1, TET2, JAK2, PPM1D,* and *SF3B1*.[45] (*DNMT3A, ASXL1, TET2,* and

JAK2 are also commonly mutated in myeloproliferative neoplasms (MPNs) and are covered in sections 1.10.1 and 1.10.2.) The majority of individuals with ARCH do not go on to develop any clinical features, however ARCH does increase the risk of developing myelodysplastic syndrome (MDS) and other blood cancers, as well as the increased risk of heart disease.

1.4 Molecular regulators of HSCs

1.4.1 Gene expression assays

Many approaches have been taken to try to discover what underpins stem-cell self-renewal, including extensive molecular profiling, and murine mutation characterisation. Gene expression analysis has made up a large proportion of these studies. In the cell, active genes are transcribed to produce messenger ribonucleic acid (mRNA), by measuring mRNA level a picture of the 'future proteome' can be generated, giving insight into cellular functions and processes. In malignancy, molecular profiling allows the effects of specific mutations to be determined and mechanisms of action, leading to disease, to be hypothesised.

The first widely used method of studying gene expression of multiple genes simultaneously was gene expression microarrays. Microarray chips have a number of microscopic spots containing nucleic acid probes specific to the mRNAs of interest. To determine the presence of specific mRNAs of interest, sample mRNAs are extracted, converted to complementary deoxyribonucleic acid (cDNA) by reverse transcription, and fluorescently labelled, when they hybridise to the target spots they fluoresce which can be detected by an imaging step that measures relative signal to noise. However, this technology lacks the ability to be reliably scaled to the single cell level, it is only suitable for use on bulk samples. For a more sensitive and specific approach, researchers turned to multiplexed quantitative Polymerase Chain Reaction (qPCR). By using targeted amplification of specific genes smaller starting amounts of mRNA can be used.

The development of microfluidic technology allowed the miniaturisation of the reaction to nL volumes to generate gene expression profiles of a set number of target genes in small cell numbers (including at the single cell level). Generally, these technologies use

sequence specific reporter probes which have a fluorophore bound at one end and quencher at the other. While the quencher is in close proximity the fluorophore will not fluoresce, but when polymerisation of the new strand degrades the probe and liberates the fluorophore from the quencher, it does fluoresce. This fluorescence can be measured by the real-time PCR machine and related to the quantity of mRNA in the sample. The most commonly used commercial technologies for these applications is the Fluidigm integrated fluidics chip (IFC) and the Biomark analyser.

A more recently developed method for assessing gene expression is RNA sequencing (RNAseq). As with microarrays cDNAs are generated from the sample mRNAs, but instead of hybridising to probes, RNAseq uses next generation sequencing to sequence the nucleic acids of the cDNAs rather than look for specific genes of interest. The benefits of RNAseq over microarray and qPCR-based approaches include not biasing sampling by looking for specific genes and being able to detect subtle changes in gene sequences like single nucleotide polymorphisms, gene fusions and splice variants. However, when looking for expression of specific genes with low expression levels, for example transcription factors, qPCR based systems are more sensitive as RNAseq approaches often have high dropout rates for genes with low levels of expression.[48] qPCR also allows the analysing of a greater number of biological samples at a substantially lower cost, although recent advances in droplet sequencing technology (e.g. 10X genomics) has lowered the costs of RNAseq even further, although this comes at the expense of sequencing depth and transcriptome coverage per cell.

As qPCR and RNAseq both measure mRNA presence to report gene expression, they share the same disadvantages. As the cellular transcriptome is continually changing, and RNA is taken at a snapshot in time, temporal events can be missed, leaving a 2D image of gene expression. Additionally, not all mRNA is translated to protein, and it is impossible to determine which mRNAs would meet alternate fates e.g. degradation. It is also important when examining gene expression data to correlate findings with functional data, i.e. do the functional properties of the cell differ with varying gene expression? By combining gene expression with functional assessment of mutant HSCs the molecular mechanisms driving malignancy can be resolved.

1.4.2 Key genes in HSCs

Maintaining HSC quiescence is a combination of repressing differentiation and division while also actively promoting the hibernating state (avoiding apoptosis, maintaining multilineage capacity, avoiding differentiation, drug efflux etc.). Differentiation and division are multifactorial processes that require precise spatial and temporal control of expression of a myriad of genes. It is therefore unsurprising that many genes associated with regulation of HSCs are involved with gene expression control, these can be broadly categorised into transcription factors, epigenetic regulators, chromatin regulators.

Transcription factors influence gene expression by directly binding and influencing the activation or repression of genes. Epigenetic regulators cause transcriptional changes by modifications to the DNA structure, for example methylation of cytosine bases, this alters the accessibility of promoters for transcription factors. Chromatin is the complex of DNA wrapped around histones. Chromatin regulators modify histones by methylation, acetylation, phosphorylation, or ubiquitination to alter the condensation of chromatin into a more accessibility open (active) or closed (inactive) state. Histone modifications affect interactions between histone subunits and with non-histone proteins. Generally, methylation is associated with a closed state while acetylation is associated with a more active state.[49]

Given the complex and interconnected network of genes involved in maintaining HSC self-renewal it is impossible to catalogue all of the genes involved, but the summary table below (Table 1) gives an overview of 45 key self-renewal genes with strong functional data or association with malignancy.

Gene	Function	Family/complex/ interactions	Effect on HSC	Links to haem cancer	ref
Bmi1	Regulate proliferation	Polycomb repressive complex 1	Loss of BMI1 induces a defect in self-renewal, while over-expression promotes self-renewal.	Overexpression correlates with poor prognosis in MDS, AML, CML, ALL, CLL, & mantle cell lymphoma	[50–52]
Bptf	Chromatin remodelling	Nucleosome-remodelling factor subunit	Knock-out mice show that BPTF is essential for maintaining the HSPC populations and their repopulating capacity	NUP98-BPTF Gene Fusion in acute megakaryoblastic leukaemia	[53,54]
Cbfa2t3h (encodes ETO2)	Transcriptional repressor and activator	Corepressor in complex with TAL1	Enhances the expression of self-renewal and survival genes while reducing the expression of growth arrest and pro-inflammatory genes	AML1-ETO fusion in AML	[55,56]
Cbx7	Chromatin remodelling	Polycomb repressive complex 1	Overexpression increases multi-lineage long term engraftment. Repression in AML cells results in terminal differentiation	Upregulated in AML	[57]
Cdkn2a (encodes $p16^{Ink4a}$ and $p19^{Arf}$)	Mediates cell cycle arrest and apoptosis	Retinoblastoma p53 pathways. Expression repressed by BMI1.	Individually deletion of $p19^{Arf}$ or $p16^{Ink4a}$ does not alter HSC engraftment or repopulation when transplanted, but triple knock-out of $p16^{Ink4a}$, $p19^{Arf}$ and p53 have a 10-fold increase in cells able to reconstitute the blood long term.	Common in non-haem cancers including skin, lung and colorectal	[58–60]
Csf1r	Colony stimulating factor 1 receptor (aka M-CSF receptor, and CD115)	Cytokine receptor, platelet derived growth factor family	M-CSF, a myeloid cytokine released during infection and inflammation, can directly induce the myeloid master regulator PU.1 and instruct myeloid cell fate change in HSC, independently of selective survival or proliferation	Therapeutic target in AML	[61–63]
Dnmt3a	DNA methylation	*de novo* DNA methyltransferases	Loss of DNMT3A increases self-renewal, HSCs can be serially transplanted up to 12 times (normal 4), differentiation block	Mutated in AML, MPN, MDS	[64,65]
Egfl7	Interacts with β3integrin to activate the AKT pathway	ECM protein	EGFL7 induces HSCs to enter cell cycle and undergo myelo-megakaryocytic differentiation	Increased expression in AML	[66,67]
Erg	TF	ETS family of TFs	Loss or ERG accelerates differentiation, leading to depletion of HSCs	Frequently overexpressed (and associated with poor outcome) in AML and T-ALL	[68]
Ezh2	Histone methylation and deacetylation	PRC2 complex	Overexpression causes increased HSC self-renewal (serial transplantation of does not exhaust)	Mutated in AML, MDS, MPN	[69]
Fli1	TF	ETS family of TFs.	Essential in haematopoietic development. FLI1 knock-out BM transplantation recipients have reduced chimerism compared with WT.	High expression in AML is an adverse prognostic factor.	[70–72]
Foxo3a	TF	Downstream of the PTEN/PI3K/AKT pathway	FOXO3A knock-out mice have reduced numbers of HSC, these HSCs are less competitive by long-term transplantation.	High expression associated with poor prognosis in AML	[73,74]

Gata1	TF	GATA family of TFs	Overexpression causes loss of self-renewal in HSCs and pushes erythroid/megakaryocyte differentiation. Not normally expressed in HSCs.	Mutations in Diamond-Blackfan anemia, acute megakaryoblastic leukaemia, transient myeloproliferative disorder	[75,76]
Gata2	TF	GATA family of TFs.	Required for HSC generation in the AGM and continues to be required for HSC survival.	Mutations in MDS, AML, and in blast crisis transformation of CML	[76,77]
Gata3	TF	GATA family of TFs	Required for the maintenance of a normal number of HSCs and for their entry into the cell cycle required for T-cell development	GATA3 variants enhance susceptibility to ALL	[78,79]
Gfi1	Recruitment of histone deacetylases and (de)methylases	Zinc finger protein	Represses the expression of genes implicated in cell survival, proliferation and differentiation. Restrains proliferation of HSCs and thereby regulates their self-renewal. GFI1 knock-out mouse has reduced number of HSCs and impaired competitiveness when transplanted.	Low expression is associated with poor prognosis in AML patients	[80–82]
Gfi1b	Recruitment of histone deacetylases and (de)methylases	Zinc finger protein	Essential for megakaryopoiesis and erythropoiesis Loss of Gfi1b increases the frequency of HSCs in BM, still able to self-renew but no longer quiescent	Role in CML (*Gfi1b* silencing inhibits imatinib-induced apoptosis). High expression in CML, AML and MPNs.	[81,83,84]
Hhex	Loss of HHEX leads to expression of *Cdkn2a*	Interacts with PRC2	Dispensable for maintenance of HSCs and myeloid lineages but essential for the commitment of common lymphoid progenitors to lymphoid lineages	Overexpressed in AML	[85,86]
Hoxa5	TF	Hox family	Involved in cell cycle control. Overexpression pushes erythroid differentiation and reduction in HSC population size. The remaining HSCs still have self-renewal and multilineage differentiation capacity.	Suppressed in AML	[87]
Hoxa9	TF	Hox family. Interacts with MEIS1 and PBX1	Over expression of *HOXA9* leads to expansion of HSCs and early progenitors, leading to myeloproliferative phenotypes in mice. Cannot progress to AML without increased expression of cofactors e.g. MEIS1/PBX1	Overexpressed in over 50% of AMLs and is highly associated with poor prognosis. MLL fusions	[88]
Hoxb4	TF	Hox family	Strong positive regulator of HSC self-renewal. Overexpression leads to HSC population expansion.	Higher *HOXB4* expression linked with better prognosis in AML	[89,90]
Ikzf1	TF	Zinc finger	Involved in the early lymphoid development and in governing the developmental pathway of lymphoid or myeloid lineage from multipotent progenitors	Mutated in ALL and CML	[91]
Itga2b (aka CD41)	Integrin	subunit of receptor complex integrin αIIbβ3	Up regulated during megakaryopoiesis. CD41 has been reported to mark myeloid and megakaryocyte progenitors within the HSC compartment.	Mutated in blood clotting disorder Glanzmann thrombasthenia	[92–94]
Kit (aka c-Kit, CD117)	Cell surface cytokine receptor tyrosine kinase	SCF receptor	Cell surface marker for HSCs. Mutant (W41) have reduced HSC competitiveness and number	Activation mutations are associated with sporadic adult human mastocytosis and with human gastrointestinal stromal tumours	[95,96]

Lmo2	Transcriptional co factor	TAL1/LMO2/LDB1 complex. Also bind LYL1.	Essential role in HSC development. *Lmo2*$^{-/-}$ mice die at E10.5 due to the lack of erythropoiesis.	Frequently deregulated gene in human T-ALL, overexpressed caused by diverse chromosomal rearrangements and other transcriptional mechanisms	[97]
Lyl1	TF	Basic-helix-loop-helix TFs. Like TAL1. Binds LMO1 and 2 and E2A	LYL1 is essential for maintaining normal HSC function in the absence of SCL.	Translocation reported in T-ALL. Also required for LMO2-induced T-cell leukaemia	[98]
Mecom (MDS1 and EVI1 complex locus)	Evi1 zinc finger proteins	Acts via PU.1 to push HSPCs towards the myeloid lineage	EVI1 knock-out results in the absence of functional haematopoietic precursors in embryos. ME-deficient mice lack quiescent, long-term repopulating cells. EVI1 overexpression expands HSC population, suppresses erythropoiesis and lymphopoiesis, and creates a myeloid-skewed phenotype.	MECOM rearrangements in AML and MDS	[99,100]
Meis1	Hox cofactor	Interacts with PBX1 and HOXA9	Required for the maintenance of haematopoiesis under stress and over the long term. MEIS1 knock-out mice contained significantly fewer numbers of long-term HSCs, which exhibit loss of quiescence. MEIS1 knock-out HSCs have reduced colony formation capacity.	Transcriptional target of MLL-fusion proteins and in cytogenetically normal AML *MEIS1* expression is associated with poor prognosis.	[101–103]
Mitf	TF	Basic-helix-loop-helix TF	Controls expression of *c-Kit* (knock-out loss of *c-Kit* expression). Involved in mast cell differentiation.	Highly expressed in systemic mastocytosis patients with *C-KIT* mutations	[104,105]
Mpl	Cell surface cytokine receptor	THPO receptor	MPL is a key regulator of megakaryopoiesis and platelet production. Essential role in the early steps of adult haematopoiesis, MPL knock-out are unable to compete with normal marrow for long-term haematopoietic repopulation	Mutated in MPNs	[106]
Myb	TF	MYB family	In mice knock-out of cMYB causes depletion of HSC pool & reduced self-renewal. Also required for thymocyte and B-cell development, myelopoiesis and erythropoiesis.	Translocations in Infant acute basophilic leukaemia, and T-ALL. Overexpression in AML, ALL, and T-ALL	[107,108]
Nfe2	TF	nuclear factor erythroid-derived 2 (NF-E2) (Erythroid transcription activator), composed of a heterodimer of NFE2 and MAFK	Involved in megakaryocyte and platelet production. Knock-down of NF-E2 in HSPCs reduces the formation of megakaryocytes, also impairs HSC activity, decreasing engraftment. In mice over expression of NF-E2 *in vivo* causes an MPN phenotype	Truncation mutations reported in MPNs although rare (~2% of cases)	[109,110]
Pbx1	Hox cofactor	Interacts with MEIS1 and HOXA9. alters HOX-DNA binding and therefore alters transcription	Conditional inactivation of PBX1 in the haematopoietic compartment results in a progressive loss of HSCs that is associated with concomitant reduction in their quiescence, leading to a defect in the maintenance of self-renewal as assessed by serial transplantation.	Chromosomal translocation causing constitutive expression of *PBX1* is associated with pre-B-cell ALL	[111]

Prdm16	TF	Repressed by HOXB4 but upregulated by HOXA9 and HOXA10	Critical for the establishment and maintenance of the HSC pool during development and after transplantation. PRDM16 deletion enhances apoptosis and cycling of HSCs.	Predictive marker for poor prognosis in paediatric AML	[112,113]
Procr (encodes EPCR)	Receptor	Endothelial protein C receptor	Well established HSC marker. A role has been suggested for EPCR in HSC homing to the BM	None	[18,114]
Runx1 (aka AML1)	TF	Interacts with PU.1, CEBPA, GATA1, GATA2, PAX5, FLI1, and ETS1	Essential in HSC development. RUNX1 deficiency decreases both apoptosis and proliferation, conflicting reports on impacts the frequency of HSCs – some say minimally effects, others increase, others decrease. Cell cycle and p53 pathways that are dysregulated in RUNX1 deficient HSCs.	Point mutations reported in AML, MDS and chronic myelomonocytic leukaemia. Translocations also common; *ETV6-RUNX1* is the most common fusion gene in childhood ALL, *AML1-ETO* one of most common fusions in AML	[115–117]
Sfpi1 (encodes PU.1)	TF	ETS family of TF	PU.1 acts by inhibiting genes that promote cell proliferation and activates those encoding cell-cycle inhibitors. Thus PU.1 is required to maintain HSC quiescence and prevent exhaustion of the HSC pool.	Reduced expression in AML. PU.1 loss-of-function heterozygous mutations or deletions found in AMLs.	[118,119]
Sh2b3 (aka LNK)	Adaptor protein	Negative regulator of JAK2	In mouse loss of LNK causes increased HSC number and self-renewal capacity in HSCs. In human expression is upregulated in $CD34^+$ cells and platelets of MPN patients, especially in JAK2 V617F-positive cases	Mutated in 5 to 7% MPN	[120–122]
Smarcc1 (aka BAF155)	SWI/SNF complex subunit	nucleosome remodelling complex	Increased HSC reconstitution activity (primary transplantation 16 weeks) with retroviral overexpression of *Smarcc1*.	Mutated in AML and malignant lymphoma	[123,124]
Smarcc2 (aka BAF170)	SWI/SNF complex subunit	nucleosome remodelling complex	Higher expression in the ESLAM fraction versus the lineage negative and $CD45^+EPCR^+CD48^-CD150^-$ fractions of BM and FL blood cells.	Not mutated in AML, CLL or malignant lymphoma	[24,123]
Tal1 (aka SCL)	TF	Basic-helix-loop-helix TFs	TAL1 is essential for the formation but not maintenance of adult HSCs. TAL1 null adult HSCs have impaired short-term repopulating ability, predominantly of the myeloid lineage, but no defects in long-term repopulation or self-renewal.	*TAL1* is ectopically overexpressed in 40–60% of T-ALL cases via chromosomal translocation, intrachromosomal rearrangement, or a mutation in the enhancer region	[98,125]
Tcf7 (aka TCF1)	TF, also chromatin remodelling	TCF/LEF family of transcription factors	TCF7 down regulated in switch from $CD34^+$ to $CD34^-$ in human, involved in self-renewal and differentiation. RUNX1 and TCF7 bind each other's promotor regions, TCF7 is necessary for the production of the short isoforms of RUNX1.	Knock-out mice develop T-cell lymphomas, but no patient mutations	[126,127]
Tet2	DNA methylation	TET family	Loss of TET2 increases HSC self-renewal as demonstrated by serial transplantation from TET2 knock-out mice.	Loss of function mutations in MPN, MDS, CMML, AML	[128]

Trib3	Pseudokinase	Endoplasmic reticulum stress modulator	*Trib3* expression is high in the HSC population. TRIB3 knock-out mice have abnormal differentiation towards mast cells. TRIB3 knock-out HSCs express mast cell genes such as *Cpa3* prematurely, poor long-term engraftment in transplantation.	*TRIB3* expression associates positively with acute promyelocytic leukaemia disease progression	[129–131]
Vwf	Glycoprotein	Mediates platelet adhesion to damaged blood vessels, and platelet aggregation.	*Vwf* expression can be used to stratify the HSC compartment by lineage priming; platelet-biased and platelet-myeloid-biased HSCs are enriched in the VWF^{+} subset. Elevated VWF expression is associated with HSCs possessing durable self-renewal potential.	Mutations cause von Willebrand disease	[24,132]

Table 1 Overview of 45 key self-renewal genes, their roles in HSCs and links with haematological malignancy

Maintenance of self-renewal involves the correct expression of many genes, some positively and some negatively regulating self-renewal capacity. Many of these genes listed in this table have strong functional data that associates them with self-renewal, and/or have been shown through malignancy to play a key role in HSCs. ALL; Acute lymphoblastic leukaemia, AML; Acute myeloid leukaemia, CLL; chronic lymphocytic leukaemia, CML; Chronic myelogenous leukaemia, CMML; Chronic myelomonocytic leukaemia, ECM; Extracellular matrix, FL; Foetal liver, MDS; Myelodysplastic syndromes, M-CSF; Macrophage colony-stimulating factor, MPN; Myeloproliferative neoplasms, PRC2; Polycomb repressive complex 2, T-ALL; T-cell acute lymphoblastic leukaemia, TF; transcription factor, THPO; Thrombopoietin

1.5 *In vitro* assays for detecting HSC potential

The experimental assessment of the functional capacity of HSCs is crucial for understanding this cell type, and how it is affected by alterations, for example mutations or aging. A number of *in vitro* and *in vivo* assays have been developed to assess HSC functional characteristics, the most common of which are discussed below.

1.5.1 CFC

HSCs can be grown clonally *in vitro*, and this was initially established in 1966 by Bradley and Metcalf, who plated BM cells in agar on feeder layers.[133] Now the colony forming cell (CFC) assay more commonly involves plating cell populations of cells in semi-solid methylcellulose-based media, the viscosity of the media allows cells to be spatially separated and grow as clonal units. CFC assays permit differentiation, and the cell types produced in the assay are directed by the cytokines present in the media. Multilineage clones can be grown, which can be assessed visually or by flow cytometry, CFU assays determine the myeloid, erythroid or megakaryocytic differentiation (but generally not lymphoid without specialised co-culture systems[134]). Both stem and progenitor cells are able to form colonies in a CFC assay as it does not measure self-renewal but proliferative potential. Serial re-plating CFCs can be used to allude to self-renewal potential of HSCs, for this CFCs are collected and replated weekly, wild-type BM can sustain (on average) three re-platings, while mutations which enhance HSC self-renewal (e.g. TET2 [135] and DNMT3A [136]) have been shown to be able to perform more serial re-platings.

1.5.2 Liquid culture

Individual HSCs can also be grown in liquid culture assays. In these assays single cells are sorted into, and cultured in, individual wells in medium supplemented with cytokines, this allows clones to be grown and the lineage output assessed by flow cytometry. As in CFCs, the cell types produced will be influenced by the cytokines added to the medium. Single cell liquid culture permits assessment of multilineage potential, cell cycle kinetics and proliferative capacity, however it does not assess long-term self-renewal, as HSCs and progenitors can both form colonies. Single cell liquid culture has the benefit over CFC assays of ensuring that resultant clones are indeed clonal; since in CFC assay colonies can overlap due to high plating density and colony picking can result in clones being mixed. In single cell liquid culture, cells are plated

in separate wells ensuring clones cannot be mixed. Single cell liquid culture has the additional benefit that the functional output can be linked with index sorting data, which allows coupling of functional output with cell surface marker expression.

1.5.3 CAFC and LTC-IC

The cobblestone area–forming cell (CAFC) assay involves culturing HSCs on a pre-established stromal layer. Unlike CFC and liquid culture which primarily assess proliferative potential, CAFCs are an *in vitro* assay for self-renewal potential. In CAFCs test cells are cultured on a stromal cell line for more than 5 weeks and assessed for the presence of haematopoietic colonies with 'cobblestone' morphology.[137] HSCs with less durable self-renewal tend to form colonies at early times while HSCs with more durable self-renewal from colonies at later time points.[37] Like the CAFC assay, the long-term culture-initiating cell (LTC-IC) assay involves growing HSCs on stromal cell layers for >5 weeks, and cells produced after this period cells are harvested and plated in CFCs. The logic is that progenitors from the original test suspension will not survive for 5 weeks before re-plating and thus the CFUs must be the progeny of HSCs. Neither the CAFC or LTC-IC are very commonly used due to disputes around their effectiveness/relatability, and high inter-lab variability.[138]

The influence of media components and additional cytokines is a feature of all *in vitro* assays – it is impossible to completely mimic the *in vivo* niche, and so the full potential of the cells can be under-reported. Consequently *in vivo* assessment of HSCs is often performed to formally demonstrate the full potential of a particular cell.

1.6 *In vivo* assays of HSC potential

1.6.1 *In vivo* multilineage assays

Both multilineage potential and self-renewal of HSC can be assessed in *in vivo* assays. Like *in vitro* assays, in order to resolve whether an individual HSC has multilineage output capacity it must be spatially separated from other HSCs – bulk assays cannot determine multilineage output.

As mentioned earlier (section 1.2), the first stem cell assay was the transplantation of BM cells which formed colonies (nodules) within the spleen of recipient animals,[8] these

nodules were of single cell origin (as determined by karyotyping studies post low dose radiation[9]) and some colonies contained multiple blood lineages – proving for the first time the multilineage capacity of HSCs on the single cell level.

Since the popularisation of single cell transplantation, it is now more common to assess the cellular output of individual HSCs *in vivo* by single cell transplantation into myeloablated recipients (alongside helper cells). The benefit of transplantation compared with *in vitro* assays is that the assay is not limited in cell type output by addition of cytokines – the clone will be able to make all cell types it can. The downside of transplantation is that it is a fairly lengthy assay and the cost and ethical considerations associated with keeping large numbers of animals is restrictive. Alternatively, HSC output can be tracked by genetic barcoding, this allows many HSCs to be assessed in the same mouse, reducing the number of recipient mice required, and permitting the studying of HSC clonal dynamics. HSCs can be labelled with a genetic barcode either by *ex vivo* transduction with a lentiviral library,[139] or by *in vivo* recombinase-induced or transposon-mediated shuffling of genetic sequences to generate unique genomic sequences in each HSC.[140] HSC output can be assessed by determining the barcode abundance within cell populations from the blood or BM.

1.6.2 *In vivo* self-renewal assays

CFU-S assays can be used to assess self-renewal, as well as multi-lineage potential, by harvesting and re-transplanting the colonies. In this way the ability of the stem cells to produce more equally potent stem cells able of sustained engraftment and blood production can be assessed. However, this assay is now uncommon, with most groups preferring serial BM or sorted cell transplantation assays. Currently, serial transplantation assays represent the gold standard to formally demonstrate HSC function by assessing their long-term multi-lineage reconstitution capacity (>16 weeks) in primary recipients and their self-renewal ability in secondary recipients.[138] When undertaking a secondary transplantation, a number of considerations must be made;

Whole BM or sorted population?

Transplantation of whole BM provides the recipient mouse with terminally differentiated cells and progenitors, in addition to the more primitive HSCs. These additional mature cells protect the recipient mouse from irradiation complications. In

order to enumerate HSCs in a test cell suspension, different doses of BM cells can be transplanted into recipient mice and the frequency of HSCs estimated using limiting dilution analysis.[1] However, a sub-optimal whole BM transplantation can be the result of the donor having a reduced number of HSCs or if the HSCs have reduced function.

Flow cytometric cell sorting has led to the isolation of purified populations of HSCs – allowing HSCs from two different conditions to be directly compared. However, flow cytometric sorting strategies are not completely generalisable across developmental stages, activation states (i.e., post-transplantation), or in mutant animals, and caution must be exercised when interpreting phenotypic populations in the absence of functional assays. A nice example of this issue is outlined by Santaguida et al. (2009)[141] where a reconstitution defect was observed upon transplantation of whole BM or HSPCs from a *Jun*B-deficient mouse model. When the group went on to look at a more purified population of HSCs they found that these cells had enhanced repopulating capacity compared to control cells. The reconstitution defects observed in whole BM and HSPC fractions resulted from a dilution of HSCs as a result of progenitor expansion.

Another example requiring care is the interpretation of transplantation data from sorted populations as evidenced by Lundberg et al. (2014)[142]. Here, a purified population of phenotypic HSCs from their JAK2 V617F mutant mouse were transplanted, but the same increase in HSC activity that was observed in whole BM transplantations was not seen. When subsets of the progenitor compartments were assessed in transplantation assays, a percentage of cells expressing lineage markers – populations normally not considered to contain HSCs - were able to initiate a sustainable disease. This study highlights a key challenge when transplanting purified populations of cells – the cells you wish to study may not be present in your purified population. When assessing purified HSCs from genetically modified mouse models, it is important to undertake new functional assays if HSCs are to be studied since a genetic mutation may affect the cell surface phenotype.

Overall, therefore, depending on the research question being asked, it may be necessary to undertake new baseline functional assays in settings where haematopoiesis has been disrupted (e.g., genetic models of malignancy). This becomes especially important

when comparing the molecular profiles of cells at the single cell level. Neither the whole BM nor sorted population strategy is perfect; the challenge is to select the most appropriate approach for the question being investigated.

To compete or not to compete?

Transplantation can be competitive or non-competitive. A non-competitive transplantation assesses whether a population containing HSCs, transplanted alone, has the capacity to engraft in an irradiated host. It can also be useful for demonstrating whether any haematopoietic phenotypes are present in the test cell suspension (e.g., a genetic mutation that results in increased red cell production)

Competitive transplant, by contrast, requires test and wild type (WT) HSCs to compete against each other to give sustained engraftment. This more accurately measures a test cell suspension's ability to outcompete normal healthy WT cells (a key property of cancer). Competitive transplantation is semi-quantitative when transplanted alongside known numbers of cells with an established repopulation capacity (e.g., 500,000 whole BM cells). Moreover, the competitor cells act as an internal technical control, controlling for such variables as the differential impact of irradiation across recipients or the technical success of an injection procedure.

1.7 Functional Heterogeneity of HSCs

The existence of heterogeneity in HSCs has been well documented from a functional standpoint. From the first CFU-S stem cell assays the cellular heterogeneity of HSCs has been alluded to; showing variability in the output of haematopoietic progenitors, variation in clone size, composition (the numbers and types of mature and primitive cells), time to detection and the number of secondary colonies they could give rise to.[9] These differences were the first hint that heterogeneity was a feature of mammalian blood stem cells. Key experiments in the 1980s tracking HSCs using retroviral inserts as genetic marks for their progeny, provided the first formal evidence of variations in self-renewal durability and also reaffirmed differences in mature cell production, prompting the eventual description of lineage-biased HSCs.[143–146]

As mentioned earlier, one of the greatest advances in deciphering the functional heterogeneity of HSCs was the introduction of multi-parameter florescence activated cell sorting for HSC purification,[16] which lead to single HSC sorting and transplantation assays.[10] Single cell transplantation provided clear evidence of the heterogeneity of HSCs, in their lineage output, longevity of graft, and graft size. However, single cell transplantations have two key limitations. The first of these is that it assesses what an HSC *can* do in a stressed environment, as opposed to what it *does* do under homeostatic conditions. The second is that it does not test how HSCs interact as a population to maintain the balance of cellular outputs. These issues are at least partially overcome by lenti- or retro-viral genetic barcoding studies where bulk HSPCs can be transplanted together and tracked individually.[147] However, these assays are still transplantation-based and are further limited by needing an *in vitro* culture period, meaning that the culture itself (or integration site of the barcode) could add bias to the population. Moreover, sequence detection limits could also add bias, as it cannot conclusively determine absence of contribution or disappearance of a clone in longitudinal studies. More advanced endogenous barcoding approaches have now been undertaken where HSCs are traced using reporter mice[140] or genetic recombination is used to mark cells *in vivo* without any additional manipulation.[42,148]

1.8 Molecular heterogeneity of HSCs

Single cell molecular profiling of HSC and progenitors has already revealed a number of findings that were previously unattainable using bulk populations. Unbiased approaches such as RNAseq have identified new genes (and associated pathways) involved in stem cell function and lineage commitment.[149,150] These same techniques have unveiled the heterogeneity within various progenitor compartments,[150] and confirmed that the molecular process by which stem cells differentiate is a gradual process (i.e., a continuum) rather than a stepwise progression though progressively more differentiated progenitors.[151] Importantly, studies that have profiled HSCs all converge on the description of significant heterogeneity within the population, but the utility of such a finding is unclear.

An important factor when considering "heterogeneity" is the difference between informative heterogeneity (biologically driven variations that result in functional differences), and generic heterogeneity (differences driven by processes occurring in all cell types e.g., cell cycle).[152] Another consideration is the extent of heterogeneity that is actually driven by technical noise: if it were possible to sequence the same cell many times, there would be technical drop-out of genes and this would differ between experiments, adding an artificial "heterogeneity" to the population. Finally, there is yet another type of heterogeneity hidden amongst molecular studies of purified HSCs – that of non-HSC contaminants. Unlike transplantation assays, where non-HSCs do not read out and are therefore not able to confuse the description of heterogeneity, molecular assays will generate 100 gene expression profiles from 100 cells. When populations are of a low purity, for example 5 or 10%, and single cell molecular studies are undertaken, it becomes incredibly difficult to assign a specific molecular

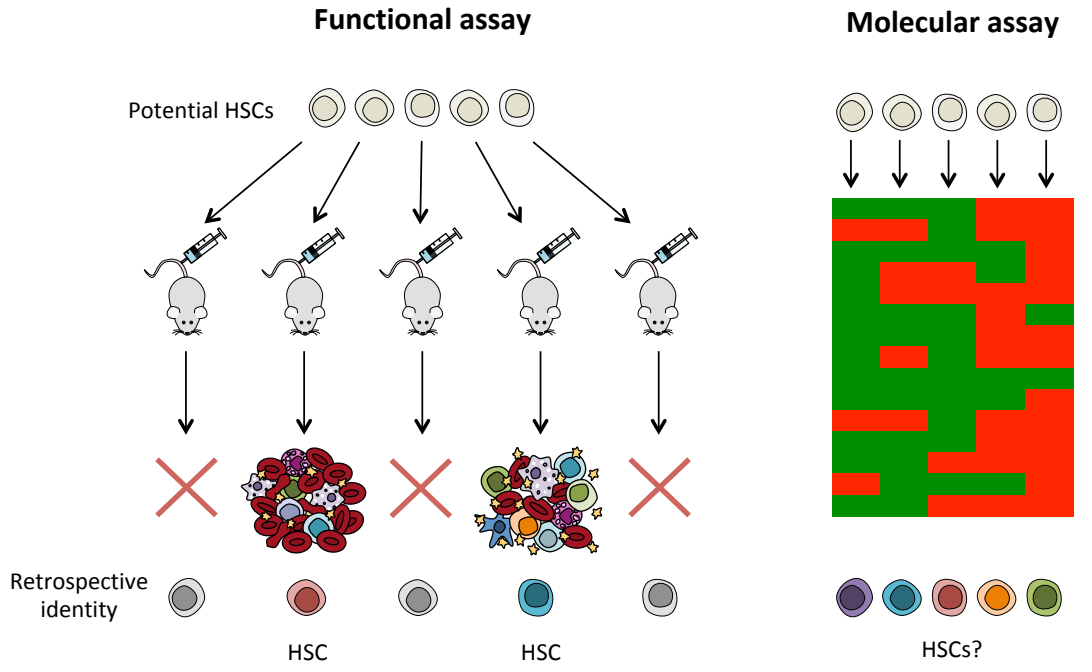


Figure 2 Selectivity of functional assays reduces experimental noise in the description of HSC heterogeneity
Depicted on the left is a single cell functional assay where 5 input cells are each transplanted into a single recipient mouse and analysed for the presence of an HSC. In this case, 3 cells do not read out and a "red" HSC subtype and a "blue" HSC subtype are easily compared. On the right are a similar set of 5 input cells where each cell is subjected to single cell gene expression profiling – all 5 cells generate a molecular profile that is unique, but it is unclear which profiles belong to HSCs and which belong to non-HSCs, making subtype classification more difficult.

programme to a specific cell function. The reality in such cases is that 90-95% of the molecular programmes are not the cell of interest and, unlike long term cell biological assays where non-HSCs are no longer represented, these represent a significant confounder for single cell molecular biology studies (Figure 2).

One of the main reasons for studying HSCs is to understand how single cells are mis-directed to drive a leukaemia. The study of 'normal' HSCs sets the stage for how 'normal' is subverted to drive malignancy and we are now in a position technically where we can start to figure that out how the molecular mechanisms of HSCs are altered, at the single cell level, by oncogenic mutations.

1.9 Myeloproliferative Neoplasms

The myeloproliferative neoplasms (MPNs) are a set of clonal, stem cell derived diseases that cause overproduction of mature myeloid cells.[153] The most well-known mutation in MPNs is the BCR-ABL fusion (first discovered in 1960[154]), CML is defined by the BCR-ABL fusion; detectable in 100% of CML patients (occasionally patients have CML symptoms without BCR-ABL, this represents a separate disease, atypical CML).[155]

The BCR-ABL fusion has been well characterised and an understanding of the mutation and molecular consequences has led to development of a potent targeted therapy in the form of Tyrosine Kinase Inhibitors (TKIs), the first of which was imatinib. Prior to TKIs, CML life expectancy was a few years (less than 20% 8 year survival[156]) and the only treatment was allogeneic stem cell transplant, since the introduction of TKIs the life expectancy for CML patients is essentially normal. TKIs are the poster child for targeted therapy; they have revolutionised treatment of this previously fatal disease. BCR-ABL and its targeted therapy highlights the power of having a comprehensive understanding of a genetic mutation and is a goal for the field in BCR-ABL-negative chronic MPNs. There are three BCR-ABL negative MPNs; polycythaemia vera, essential thrombocythaemia, and myelofibrosis.[157]

1.9.1 Polycythaemia vera

Polycythaemia vera (PV) is characterised by an overproduction of red blood cells and is diagnosed by an increased erythrocyte volume; haemoglobin level greater than 16.0g/dL in women or 16.5 g/dL in men, or haematocrit greater than 48 percent in women or greater than 49 percent in men. Low erythropoietin (EPO) levels in combination with high red cell count is indicative of PV, as it suggests cytokine independent activation of the JAK/STAT signalling pathway. In almost all PV patients this activation is caused by JAK2 mutations, either V617F or exon 12. The incidence of PV is approximately 0.4 - 2.8 individuals per 100,000 per year,[158] with slight male preponderance (1.2:1).[159] The median age at diagnosis is 60,[159] and the median survival 14 years.[160]

The greatest risk for PV patients is thrombosis. PV patients can be treated with venesection to reduce the number of circulating erythrocytes and thus the viscosity of the blood, thereby reducing the risk of thrombosis.[161] With the application of venesection, mean survival for PV patients increased from 2 to 10 years.[162] Patients can also be treated with low-dose aspirin to reduce the risk of thrombosis, hydroxyurea to suppress blood cell production,[163] or interferon (IFN)-α to induce differentiation in JAK2 V617F mutant HSCs.[164]

1.9.2 Essential thrombocythaemia

Essential thrombocythaemia (ET) is characterised by overproduction of platelets and diagnosed on high platelet count - above 450,000 platelets per microliter of blood (with exclusion of reactive thrombocytosis causes, such as infections, iron deficiency and other haematological or non-haematological neoplasms). The incidence rate of ET is 0.38 – 1.7 individuals per 100,000 /year, with a 2:1 female bias.[158] The median age at diagnosis is 65-70,[165] and the median survival is 20 years.[160] ET patients are at risk of thrombosis and haemorrhage as well as transformation to AML (2% risk at 15 years), and MF (9% risk at 15 years). Like PV patients the main risk for ET patients is of thrombosis. Patients can be risk stratified by age (high risk aged over 60), presence of previous thrombosis, and platelet count >1500x10^9/ul. High risk ET patients can be treated with hydroxyurea or IFN-α to reduce thrombotic risk.[166]

1.9.3 Myelofibrosis

Myelofibrosis (MF) is the most severe of the BCR-ABL negative MPNs and is characterised by fibrosis of the BM. MF can occur as the primary disease (pMF), or ET or PV can transform to MF (secondary MF, sMF). MF is diagnosed by BM morphology; collagen and/or reticulin fibrosis, with increased megakaryocyte proliferation and atypical megakaryocytes (large with atypical nuclear:cytoplasmic ratio and irregularly folded nuclei).[157] This fibrosis is thought to be the result of the MPN clone producing cytokines (e.g. fibroblast growth factor), which stimulate the growth of connective tissue in the BM. This overproduction of fibrotic tissue replaces the haematopoietic cells within the BM, leading to progressive pancytopaenia.

The incidence rate of pMF is 0.1 to 1 per 100,000 individuals per year,[158] with patients presenting at a median age of 64. The median survival is 6 years.[160] Of the BCR-ABL negative MPNs, MF has the highest risk of transformation to AML of 5-30%.[167] Currently the only curative treatment for MF is stem cell transplantation, however the risk of transplant-related death or morbidity is very high (over 50%). Other treatments, like JAK2 inhibitors, are palliative and have not been shown to be disease modifying.[157]

1.10 Mutations in MPNs

The development of a cancer is a multi-step process in which acquired genetic mutations confer an advantage to cells, allowing them to outcompete normal cells. All haematological malignancies can trace their origins to single stem or progenitor cells in the blood system.[168] In order for these clones to evolve from a single cell - and expand to the extent where clinical features become apparent - the population of cells must have a significant clonal advantage over non-mutant blood HSCs. To achieve such a clonal advantage, two functional properties must be operative: 1) increased proliferation: one cell needs to expand to billions, and 2) HSC self-renewal: the clone must be continually sustained by newly created daughter HSCs. Whether a single mutation can drive both properties, however, remains unclear.

The BCR-ABL fusion protein is defining of CML, and consequently is present in all patients,[155] most often alone.[169] Despite its prevalence, it has been suggested that this mutation alone is insufficient to drive CML, as BCR-ABL HSCs have been reported to

have a decreased probability of self-renewal divisions and a bias towards differentiation.[170] Indeed BCR-ABL has been detected in healthy individuals,[171,172] suggesting the presence of this mutation alone is insufficient to drive disease. This is supported by a mouse model of BCR-ABL lacking the capacity to repopulate a secondary mouse on serial transplantation.

Koschmieder et al. (2005)[173] generated a BCR-ABL transgenic model expressing BCR-ABL (under the control of the SCL promoter) in primitive haematopoietic cells. These animals developed a myeloproliferative disease that resembled human CML - with neutrophilia, leucocytosis and myeloid cell infiltration of extramedullary organs. Primary BM transplantation or sorted Lin^-, $Sca1^+$, $c\text{-}Kit^+$ cells (LSK; less than 10% HSCs[138]) resulted in a marked decrease in chimerism over the 5-month follow up period, indicating a failure of long-term reconstitution, this was accompanied by a lack of disease phenotype. A follow-up study[174] transplanted whole BM or LSK cells from the same BCR-ABL mouse line into a different strain of recipient (FVB/N) and were able to recreate the disease phenotype. However, cells from primary recipients failed to initiate a leukaemic phenotype in secondary mice, due to low chimerism. Moreover, leukaemic mice show a significant reduction in BM HSCs (LSK, $CD34^-$, $FLK3^-$) and increased multipotent progenitor (MPP; LSK, $CD34^+$, $FLK3^+$) number compared to controls, leading to the idea that BCR-ABL mutation might be responsible of increased differentiation of HSCs.

Since the majority of cancers are highly mutated, resulting in a large degree of genetic heterogeneity, it can be challenging to accurately assess the impact of an individual or a small number of mutations on a particular disease.[175] By contrast, MPNs are relatively simple in terms of their genetics[175] and therefore are a highly tractable set of diseases in which to study the individual contributions of a given mutation on the biology of the clone. MPNs provide a window into the very earliest stages of tumorigenesis, typically carrying fewer than 3 genetic driver mutations.[176] In this thesis I primarily focus on two of the most common genetic drivers in MPNs, the most commonly commutated pair, JAK2 V617F and TET2 loss of function, investigating their properties individually and in combination. The next section describes a wider range of mutations in MPNs.

1.10.1 Signalling mutations in MPNs

The vast majority of MPN patients have mutations in one of three genes; *JAK2, MPL* or *CALR*. Mutations in these 3 genes are nearly always mutually exclusive and all lead to hyperactivation of the JAK/STAT signalling pathway.[177] Across the 3 MPN subtypes there is a different prevalence of these 3 mutations. In PV the vast majority of patients have a JAK2 mutation (95%) with CALR and MPL mutations virtually absent (0%). In ET approximately 60% of patients have a JAK2 mutation, 20% have a CALR mutation and 3% a MPL mutation. In pMF JAK2 mutations are found in approximately 60% of patients, 25% of patients have a CALR mutations, and 7% have a MPL mutation.[157] Patients who do not have mutations in any of these 3 genes are referred to as triple-negative. Approximately 17% of ET and pMF patients, and 7% of PV patients are triple negative.[177]

Most cytokine receptors lack intrinsic kinase activity and therefore employ the Janus kinase (JAK) family of proteins to carry out phosphorylation of downstream proteins. Under normal physiological conditions, cytokines such as EPO, THPO and Granulocyte colony-stimulating factor (G-CSF) bind to their monomeric receptor (EPOR, MPL or G-CSFR), ligand binding induces receptor dimerization and

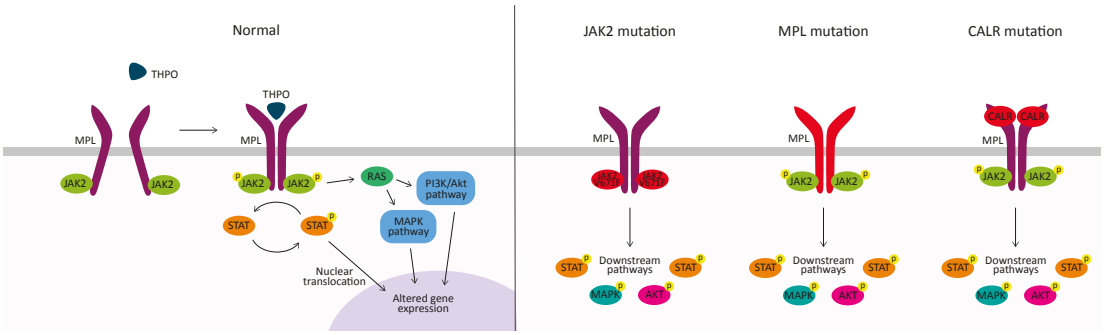


Figure 3 Activation of JAK/STAT signalling pathway under normal conditions and with MPN mutations

Under normal conditions THPO binds MPL causing a conformational change, leading to dimerization, and cross activation of intracellularly bound JAK2. This causes activation of downstream JAK/STAT signalling pathway and other signalling targets. In MPNs, when JAK2 is mutated (V617F) it is hyperactivated and does not require cytokine stimulation to induce receptor dimerization and activate downstream signalling pathways. Likewise, when MPL is mutated it dimerises without cytokine stimulation, causing inappropriate downstream signalling. When CALR is mutated it binds to the MPL extracellular domain resulting in activation of MPL and the consequent downstream pathways.

activation,[178] this dimerization brings the intracellular bound JAK2 into proximity to allows cross-phosphorylation (Figure 3). Phosphorylated JAK2 is able to phosphorylate signal transducers and activators of transcription (STAT) family transcription factors, which are then able to translocate to the nucleus and initiate transcription.[179]

JAK2

The most common mutation in BCR-ABL negative MPNs is JAK2 V617F. This mutation was first reported in 2005 [180–184] and is present in 95% in PVs, and ~50% in ET and MFs.[180] The JAK2 V617F mutation has also been reported at lower frequencies in other myeloid malignancies including CMML (6%), BCR-ABL-negative CML (19%), and MDS (3%).[182,185,186]

JAK2 V617F is a gain-of-function point mutation (a G to T substitution that causes a valine to phenylalanine switch) in the pseudokinase domain of JAK2.[187] This mutation disrupts the inhibitory function of this domain and results in a hyper-activation of JAK2 and consequently the downstream JAK/STAT signalling pathway and other signalling targets [181] (Figure 3). Most PV patients who lack a JAK2 V617F mutation have a JAK2 exon 12 mutation, exon 12 is the linker between Src homology 2 domain (SH2) and the pseudokinase domain. These mutations are mostly in-frame small insertions or deletions.

Since its discovery, many studies have been undertaken to resolve the molecular effects of JAK2 V617F as it is unclear how the same mutation can be associated with three distinct diseases. A number of theories exist to account for how the same mutation could be associated with three phenotypes. The gene-dosage hypothesis suggests the number of copies of JAK2 V617F (hetero- or homozygous) and allele burden (ratio of mutant to non-mutant cell) correlates with different disease phenotypes; homozygous clones and high allele burden are associated with PV. Homozygous clones are rarely seen in ET patients, these patients also tend to have lower allele burden.[188] Microenvironmental factors, such as serum erythropoietin levels, have also been proposed as a factor driving distinction between diseases phenotype.[189] Genetic factors have been suggested to predispose individuals to the development of MPNs and contribute to phenotypic diversity. Having a first-degree relative with an MPN has been reported to increase the risk of developing an MPN by 5-7 fold.[179] Another genetic

factor reported to increase predisposition to MPN is the 46/1 haplotype block. This haplotype block contains the JAK2 gene and increases the odds of developing an MPN by 3-4 fold.[179] As JAK2 V617F is commonly co-mutated, many groups have suggested that the co-mutated genes, and even the order they arise in,[190] contribute to MPN progressions and phenotype. It has also been suggested that JAK2 V617F alone is insufficient to drive disease.[191] This idea is supported by a study of ~4,000 individuals attending outpatient clinics carried out in China, where it was found that nearly 1% of patients have a JAK2 V617F mutation without presenting an MPN phenotype.[192]

These concepts about MPN progression are not mutually exclusive. Other theories explore the ways in which genetic factors and collaborating mutations interact. It has been suggested the 46/1 haplotype block could contribute to MPN disease progression in two ways: either the block increases occurrence of JAK2 mutation (hypermutability hypothesis) or, as the fertile ground hypothesis posits, the block provides a selective advantage to the cell once the mutation has occurred, allowing it to acquire further mutations.[193]

MPL

MPL is the receptor for the cytokine THPO. MPL has no intrinsic tyrosine kinase activity, instead JAK2 is bound to MPLs intracellular domain. When THPO binds MPL it dimerises, and a conformation change occurs resulting in rapid the phosphorylation and activation of JAK2. This then phosphorylates MPL and activates downstream signalling pathways (STATs, AKT and ERK).[194,195]

A number of activating mutations in MPL have been found in ET and pMF patients, the 2 most common locations being W515 (W to L, K or less commonly R, A, or G) and S505 (S505N). Tryptophan W515 is located at the junction between the transmembrane and cytosolic domains of the receptor and is involved in dimerization and activation.[196] While serine S505 is located in the transmembrane domain, the MPL S505N mutation stabilises the receptor in the active, dimerised, form, thus causing hyper-activation.[197]

In order to determine the role of MPL mutation on haematopoietic cells, retrovirally expression of *MPL W515L* in cell lines and BM has been performed.[198] In cytokine

supplemented conditions expression of *MPL W515L* in cell lines and BM drove no difference in proliferation, however when cultured in cytokine free conditions *MPL W515L* transduced cells exhibited cytokine independent growth. When BM transduced with *MPL W515L* was transplanted, an MF like phenotype was observed; with fibrosis of BM, splenomegaly with thrombocytosis and leukocytosis (but no change in haematocrit). *MPL W515L* expressing cells were not serially transplanted so no conclusion can be drawn about the self-renewal of these cells.[198]

CALR

Calreticulin has 2 main functions under normal conditions; 1. Maintaining calcium homeostasis, and 2. As a chaperone within the ER. In the latter, its role is to prevent misfolded proteins from being exported from the ER to the golgi apparatus.[199] CALR does not appear to be directly involved in JAK/STAT signalling, however it has been found that mutant CALR binds to, and hyper-activates, the THPO receptor MPL via its extracellular domain (not the ligand binding site), but the precise mechanism for this is currently unclear.

The two most common CALR mutation in MPN patients are a 52 base-pair deletion (commonly known as del52, c.1092_1143del) and a 5 base-pair insertion (ins5, c.1154_1155insTTGTC). Both mutations result in an exon 9 (c-terminal) frameshift mutation, causing removal of the ER retention motif KDEL (and changing the charge of the C terminal domain from negative to positive). It is hypothesised that this mutation causes a conformational change which allows the N-terminus of CALR to bind to MPL via an N-glycosylation site, causing hyper-activation and as a consequence hyperactivation of the down steam JAK/STAT signalling pathway.[200]

A mutant CALR knock-in mouse model[201] has been developed with a conditional human CALR c-terminal, with a 52 base pair deletion, induced by polyinosinic-polycytidylic acid (pIpC) treatment (Mx1-Cre). Heterozygous mice have been reported to have an ET-like phenotype (characterised by elevated platelet count), while homozygous mutant mice have a MF-like phenotype (displaying elevated platelets, splenomegaly, and reduced BM cellularity). Both homozygous and heterozygous mice have been reported to have an elevated number of phenotypic HSCs (ESLAM), but neither have a competitive advantage upon primary or secondary transplantation.

1.10.2 Mutations in epigenetic regulators

Outside of mutations causing hyper-activation of the JAK/STAT signalling pathway, the second most common 'type' of mutation in MPNs are mutations in epigenetic regulators. Epigenetic mechanisms position DNA in a state that allows expression of appropriate genes for that cell type. Disruption in epigenetic regulators can lead to an overly restrictive or permissive gene expression landscape. A restrictive epigenetic pattern could prevent turning off of proliferation or self-renewal pathways, or prevent activation of differentiation or apoptotic programmes, while a permissive landscape could allow expression of oncogenes.[202] A disrupted epigenome can also affect telomere maintenance and DNA repair pathways, influencing the rate at which mutations arise and are repaired.[203]

MPN patients commonly have a tyrosine kinase activating mutation alongside a mutation in an epigenetic regulator. The most commonly mutated epigenetic regulators in MPNs are *TET2* (12%), *ASXL1* (5%), *DNMT3A* (5%), *EZH2* (~3%), and *IDH1/2*(~1.5%) (Figure 4).[204] Mutations in these genes are also commonly found in ARCH (section 1.3.2).

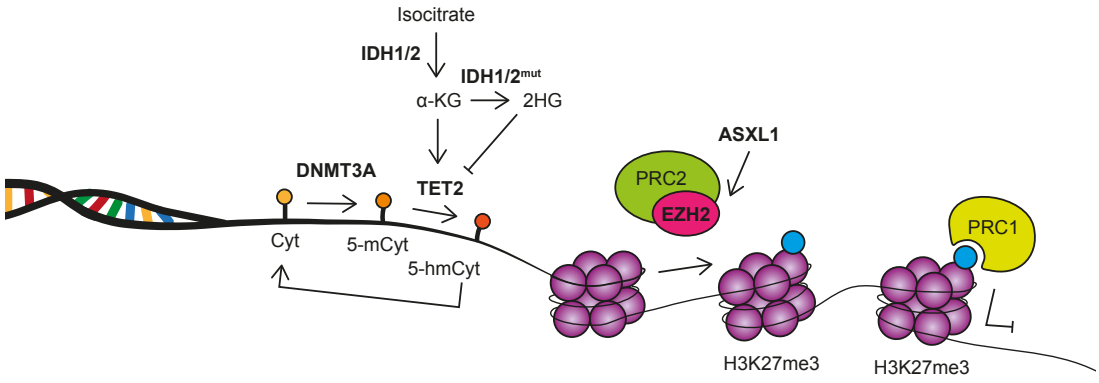


Figure 4 The normal role of commonly mutated epigenetic regulators
DNMT3A is a de novo methyltransferase; catalysing the conversion of cytosine (Cyt) to 5-methyl-cytosine (5-mCyt). TET2 in turn converts 5-mCyt to 5-hydroxymethyl-cytosine (5-hmCyt). For this TET2 requires alpha-ketoglutarate (αKG), the production of which is controlled by IDH1/2. Mutations in IDH1/2 result in 2-hydroxyglutarate (2HG) being produced instead of αKG, 2HG inhibits TET2. At the histone level, the polycomb repressive complex 2 (PRC2) trimethylates histone H3 at lysine 27 (H3K27me3), repressing gene expression, H3K27me3 can be stabilised by PRC1. EZH2 is part of the PRC2 complex and ASXL1 interacts with PRC2.

DNMT3A

Methylation of the 5-C position of cytosine is a stable epigenetic mark that acts to repress gene expression. DNA methylation is crucial for regulation of gene expression and is essential for normal development. The DNA methyltransferase (DNMT) family is responsible for the methylation of DNA. DNMT1 primarily methylates hemimethylated DNA at the replication foci during S phase of the cell cycle, thereby maintaining pre-existing DNA methylation patterns.[205] DNMT2 is a methyltransferase homolog that methylates transfer RNA (tRNA). DNMT3A and B are primarily de novo methyltransferases, establishing new methylation patterns during development and differentiation. DNMT3L has no catalytic activity but assists DNMT3A and B by increasing their ability to bind S-adenosyl-L-methionine, the methyl group donor.[205]

DNMT3A mutations are common in individuals with AML, MPN, MDS, T-ALL and ARCH.[206] The most common mutation in DNMT3A is a missense mutation at arginine R882, located in the methyltransferase domain, at this location the most common mutation is R882H (2/3rds of patients), followed by R882C (1/3rd), and occasionally R882S and R882P (3%). The R882 residue is located at the central interface of the DNMT3A tetramer where it interacts with the DNA backbone. Mutations at arginine R882 effect the function of DNMT3A by altering the flanking sequence preferences for DNA binding (altering the CpG locations which are methylated and hence gene expression).[207]

Two types of mouse models have been developed in order to elucidate the role of DNMT3A in disease and HSC function; DNMT3A knock-out models, and *Dnmt3a R878H* knock-in models. *Dnmt3a R878H* (mouse homolog to *DNMT3A R882H*) mice develop AML, and have splenomegaly and enlargement of progenitor (LSK) compartment.[208] Meanwhile loss of DNMT3a ($DNMT3A^{-/-}$) has been shown to promote self-renewal, leading an expansion in the HSCs population. This increase in self-renewal is most apparent in serial transplantation experiments; whereas WT HSCs can only sustain 4 rounds of serial transplantation before exhausting, $DNMT3A^{-/-}$ HSCs can undergo at least 12 rounds of transplantation,[64] leading the authors to report that they have been 'immortalised' *in vivo.*

TET2

The ten-eleven translocation (TET) family of proteins (TET1-3) are dehydrogenases responsible for catalysing the conversion of 5-methyl-cytosine (5-mCyt) to 5-hydroxymethyl-cytosine (5-hmCyt).[128] Unlike 5-mCyt, which is maintained during replication by DNMT1, 5-hmCyt is lost during cell division.[206] The first TET family member identified was TET1 which was found as a fusion partner of MLL in AML (a t(10;11)(q22;q23) translocation), TET2 and 3 were subsequently identified as homologues. TET1 and 3 have a CXXC and therefore bind DNA directly while TET2 binds indirectly. In haematopoietic cells TET2 is the highest expressed member of the family.[128]

Loss of function TET2 mutations are present in a range of myeloid malignancies, including MPNs, MDS, CMML, and AML, as well as being a frequent mutation in ARCH.[128] TET2 mutations occur throughout the gene (no hotspots), and are variable; small insertions, deletions, and nonsense mutations, have all been reported, the commonality between these mutations is that all induce loss of function in the protein.[209] In MPNs, *TET2* is the most commonly co-mutated gene with *JAK2 V617*. Loss of function TET2 mutations give a self-renewal advantage to HSCs as determined by numerous TET2 knock-out mouse models;[135,210–212] these models all display expansion in HSC and progenitor cell compartments and increase *in vivo* repopulation capacity. TET2 knock-out mouse models are discussed in section 3.3.2.

IDH1/2

TET2 function requires alpha-ketoglutarate (αKG), αKG is produced from isocitrate by oxidative decarboxylation, a process that is catalysed by the isocitrate dehydrogenase (IDH) family of proteins. IDH1 and 2 are the family members implicated in haematopoiesis. In MPNs, mutations in IDH1/2 alter the enzymatic activity of these proteins and catalyses the reaction of αKG to 2-hydroxyglutarate (2HG). While TET proteins require αKG, 2HG impairs TET2 function. Targeted inhibition of IDH2 can restore leukaemic cell differentiation.[128] Mutations in *IDH1* and *IDH2* have been described in AMLs (20%) and MDS (5%) patients.[213] The most common mutation sites in *IDH1* is at R132, and in *IDH2* is at R140, and R172.[213]

Unlike TET2 knock-out mouse models, which have enhanced HSC self-renewal capabilities, IDH1 mutant mice have reduced numbers of HSCs and HSC self-renewal.[214] This is due to IDH proteins' TET2 independent roles; there are over 80 other αKG dependent enzymes.[215] IDH1/2 mutations have also been shown to affect DNA damage by decreasing levels of ataxia telangiectasia mutated (ATM) (via altered methylation).[214] ATM is recruited to double strands breaks and has a role in cell cycle during DNA repair. Reduced ATM levels in IDH1 mutant animals leads to impaired DNA repair and as a consequence increased sensitivity to DNA damage and then apoptosis. This could explain the reduced number of HSCs seen in IDH1 mutant mice.[214]

EZH2

Beyond epigenetic alterations to DNA, epigenetic modifications can also affect gene expression through modification of histones. Histones are susceptible to many types of modification include acetylation, methylation, phosphorylation, ubiquitination, and sumoylation. The polycomb repressive complex 2 (PRC2) trimethylates histone H3 at lysine 27 (H3K27), one of the most common forms of histone modification. The complex is made up of 3 subunits EZH2, EED, and SUZ12. EZH2 (enhancer of zeste 2) is the catalytic subunit of PRC2. The trimethylation of H3K27 (H3K27me3) is a mark of transcriptional silencing, this mark is stabilised by PRC1, which acts directly to silence target genes. Polycomb mediated repression is considered more readily revisable that DNA methylation.

Trimethylation of histones (H3 lysine 9 (H3K9), histone H3 lysine 27 (H3K27), and histone H4 lysine 20 (H4K20)) has also been suggested as a prerequisite for DNA methylation. Components of these methylation systems interact with DNMTs, bringing DNMTs into proximity with areas for methylation. EZH2 has been shown to interact directly with DNMTs. Overexpression of *Ezh2* has been shown to increase CpG methylation, and knock-down has been shown to reduce methylation in regions known to be targets for EZH2 mediated H3K27 methylation.[205]

EZH2 has also been reported to have actions outside of its PRC2 histone methylation activity; as part of the PRC2 complex EZH2 methylates non-histone protein substrates (e.g. STAT3, GATA4, talin, and RORα), influencing transcriptional silencing and

activation.[216] EZH2 also acts independently of PRC2, as a co-activator for transcription factors, such as AR-associated complex, NF-κB signalling, TCF/β-catenin and PCNA, and β-catenin and ERα.[216] Overexpression, gain of function, and loss-of-function mutations in *EZH2* have been detected in cancers, suggesting that EZH2 can function as a tumour suppressor gene or as an oncogene.[216] In MPN and MDS patients, *EZH2* mutations generally are loss of function (numerous mutations have been detected through gene including nonsense or frameshift mutations).[217] In MPNs, loss-of-function EZH2 mutations have been found in ~10% of patients, like TET2 these mutations are associated with reduced survival and increased transformation risk when co-mutated with JAK2 V617F.[217]

ASXL1

Additional sex combs like 1 (ASXL1) is involved in chromatin remodelling though interactions with polycomb complex proteins (EZH2 and SUZ12), and therefore is involved in the placement of the histone repressive mark H3K27me3.[218] Inhibition of ASXL1 results in loss of the H3K27me3 mark. The exact role of ASXL1 in the placement of this repressive mark is not known, but it is though that it may be involved in the recruitment of the PRC2 complex to the required loci. ASXL1 has also been reported to interact with the heterochromatin repressive component HP1α. HP1α is also involved in epigenetic silencing of genes and binds histone H3.[218]

ASXL1 mutations in myeloid malignancies are most often frameshift or non-sense mutations in exon 12 of the gene.[218] These mutations result in truncation of the protein before the c-terminal plant homeodomain, which is thought to bind methylated lysines.[218] *ASXL1* mutations have been reported in MPN (8% of patients), MDS (10%), and CMML (40%).[219] Within MPNs *ASXL1* mutations are relatively frequent in pMF patients (34%) but rare in PV or ET patients.[218]

In MPN, ASXL1 and JAK2 mutations appear to be mutually exclusive. Both ASXL1 and JAK2 have been reported to interact with HP1α; JAK2 causes phosphorylation of histone H3Y41 and excludes HP1α from chromatin.[220] Activating JAK2 mutations and loss of function ASXL1 mutations, it can be hypothesised, would both exclude HP1α from chromatin. This could provide a functional link between the two mutations, but this is yet to be proven.[218]

1.10.3 Patients with only a JAK2 mutation

There still exists a substantial cohort of JAK2 V617F-positive MPN patients with no known additional driver mutation. It must then be considered that if JAK2 V617F does not give an intrinsic clonal advantage (as suggested by mouse models below, section 1.11.1), some additional factor(s) must influence a JAK2-mutant HSC outgrowth relative to the other HSCs in the body. Several possible explanations exist including: 1) additional genetic or epigenetic drivers not yet identified by exome sequencing (e.g., long non-coding RNAs, enhancer elements, etc) indeed it has been suggested that, across all cancers, half of driver mutations are yet to be discovered;[221] 2) inherited genetic risk factors where some patients are more susceptible to cells gaining a clonal advantage (e.g. SNP profiles such as the 46/1 haplotype); and 3) micro-environmental factors (the physical niche itself or secreted cytokines) that encourage the outgrowth of mutant subclones.[222] It is also interesting to note that JAK2 V617F mutations are found in a substantial percentage of older individuals with no obvious blood phenotype.[222] This latter finding would be consistent with a low number of poorly competitive JAK2 V617F HSCs giving rise to more mature cells on a per-HSC basis, but an insufficient number to create an observable phenotype in an individual.

1.11 MPN Mouse models

1.11.1 JAK2 V617F knock-in mouse models

A number of mouse models have been generated to gain understanding of the effect of common MPN mutations on disease initiation and progression. For a comprehensive review of more than 15 distinct JAK2 V617F mouse models see Li et al.[191], including retroviral over-expression, transgenic, and targeted gene knock-in models. The most physiologically relevant of these models are the genetic knock-ins where the gene is under the control of endogenous elements. While some phenotypic diversity is observed between these knock-in JAK2 V617F models, they all recapitulate the main aspects of human disease. Although all groups report myeloid expansion from their models, the effect of the JAK2 V617F mutation on the self-renewal capacity of HSCs, as determined by secondary transplantation, has not been reported on so widely. Currently there are three JAK2 V617F knock-in models for which secondary transplantation data is available, developed by groups in Boston, Paris and Cambridge.

The Boston model (Mullally et al. 2010[223])

Upon competitive primary transplantation of LSK cells from the Boston model (E2A-Cre/JAK2 V617F heterozygous) the group saw a myeloid expansion with over 80% of myeloid cells possessing the JAK2 mutation at 65 weeks post transplantation ([223] followed up in [224]), suggesting a strong mutant cell advantage. However, when BM cells from the primary LSK transplants (2:1 JAK:WT) were tested in secondary transplantation, at 16 weeks post-transplantation[164] JAK2 mutant cells did not display larger graft size than WT counterparts (Figure 5E[164]) (primary receipts shown in Figure 4D,E[223], and Figure 2A[224]). This suggests that, despite their relative advantage in primary transplantations, JAK2 mutant HSCs do not have the self-renewal advantage required for sustained clonal growth. Supporting this notion, when this mouse was crossed to a TET2-deficient mouse, the HSC self-renewal was enhanced, resulting in secondary chimerism from double mutant animals being over double of that from JAK2 mutant alone (in secondary transplantation JAK alone chimerism ~40%, double mutant ~90%).[225]

The Paris model (Hasan et al. 2013[226])

Primary transplantation of BM cells from the Paris model (a conditional Vav-Cre/JAK2 V617F heterozygous model) resulted in an expansion of the JAK2 mutant clone (Figure 4A[226]). Here, transplantation experiments were performed using whole BM which would control for any potential selection imposed by sorting stem/progenitor cells. Mice transplanted with a 30:70 mutant:WT graft resulted in mutant cells representing ~85% of the total myeloid cells, ~40% of the lymphoid and 99% of the total phenotypic LSK SLAM cells. Upon secondary transplantation (from 30% mutant: 70% WT BM transplantation primary, vehicle treated, over 75% chimerism by end of primary, Figure 6A[226]) declining JAK2 mutant graft size was observed - 45% at 4 weeks falling to ~25% at 16 weeks (Figure 6C[226]), again suggestive of fewer (or less robust) functional HSCs being produced in the primary competitive transplant.

The Cambridge model (Li et al. 2010[227] and Li et al. 2014[228])

The Cambridge model (heterozygous[227] and homozygous[228] human JAK2 V617F under endogenous promoter) shows the most pronounced effect on the relative competitiveness of JAK2 V617F HSCs. When whole BM transplantations were performed using the JAK2 V617F heterozygous pIpC inducible model, elevated

myeloid lineages were observed, but there was no striking effect on HSC donor chimerism (Figure 6C[227]). In secondary competitive transplantations, a marked self-renewal disadvantage in JAK2 V617F mutant HSCs was observed (Figure 6D[227]).

The JAK2 V617F homozygous model also showed an HSC self-renewal defect, with some indication already present in the primary whole BM transplantation, as represented by a progressive decline in chimerism (Figure 6D[228]). The earlier presentation and exhaustion of functional HSCs may be due to the homozygosity of the JAK2 mutation, the fact that this model expresses human JAK2 rather than mouse, or as a result of the comparatively lower cell dose used for transplantation (ie. Large numbers of donor cells can sustain haematopoiesis in secondary transplantations). To assess if reduced chimerism was due to a reduced number of HSCs being transplanted (Fig. 6A/B[227]), purified HSC from JAK2 V617F homozygous, heterozygous and WT donors were transplanted and confirmed reduced repopulation in recipients of homozygous cells. (Figure 6I[228]). When secondary transplantation of whole BM from the homozygous model was performed,[229] the relative loss of functional HSCs was more pronounced (Figure 4D[229]).

Taking the data from these three models together, we can conclude JAK2 V617F drives a hyperproliferative advantage, as demonstrated by large myeloid clones in primary transplantation recipients, but the long-term self-renewal capacity of HSCs is not increased compared to WT HSCs.

1.11.2 TET2 knock-out mouse models

Due to the different targeting strategies used to produce the knock-in, JAK2 V617F mouse models exhibit considerable variability in their phenotype and HSC properties. Conversely mouse models of loss of function TET2 are all very similar, as they all have complete loss of TET2 function.

Studies on models of loss of TET2, generated by Ko et al. 2011,[210] Moran-Crusio et al. 2011,[135] Li et al. 2011,[211] and Quivoron et al. 2011,[212] all report an increase in size of the LSK compartment (proportion and absolute number), increased serial re-plating ability *in vitro*, and a competitive transplantation advantage. All groups also report an abnormal skewing of blood parameter towards a CMML phenotype (Table 2).

The TET2 knock-out model used throughout this thesis is the Ko et al. (2011[210]) model, generated by disruption of TET2s catalytic domain (by deletion of exons 8-10, exon 9 contains HxD motif required for catalytic activity). These mice do not display any changes in myeloid, lymphoid, erythroid populations (by flow cytometry) but do have increased numbers of LSK and LK populations in their BM and spleens. TET2 knock-out mice and recipients of TET2 knock-out BM have increased numbers of LSKs. Transplantation of BM from TET2 knock-out animals has shown HSCs from these animals have an increased repopulating capacity.[210] Like other loss of TET2 models, this model shows that loss of TET2 leads to an increase in HSC self-renewal.

	Quivoron et al. (2011) [212]	**Moran-Cruzio et al. (2011)**[135]	**Ko et al. (2011)**[210]	**Li et al. (2011)**[211]
Models	Exon 9 (gene trap). Exons 10–11 (conditional deletion)	Exon 3 (conditional deletion)	Exons 8–10 (conditional deletion)	Tet2 disruption 6 bp upstream of the transcription start
5-hmC levels	Decreased	Decreased	Decreased	Decreased
BM Lin^-, $c\text{-}Kit^+$		Increased, serial re-plating	Increased	Increased
BM LSK $CD150^+$, $CD48^-$	Increased	Increased	Increased	Increased
BM progenitors	Increased CMP, MEP	Increased CMP	Increased CMP	Increased GMP
Extramedullary haematopoiesis	Spleen, Liver	Spleen, 20 weeks	Spleen 12 weeks	Spleen, 2–4 months
Repopulation capability *in vivo*	Increased	Increased	Increased	Increased
Diseases	CMML-like	CMML-like (70%)	Accumulation of $CD115^+$, F4/80+ cells	CMML-like and spectrum of myeloid malignancies
Disease transplantability	Yes	Yes		Yes

Table 2 Four mouse models of loss of TET2 show very similar features
Adapted from Solary et al. (2013)[128]

1.11.3 JAK2 V617F/TET knock-out double mutant models

As JAK2 V617F is the most common mutant in MPNs, and a loss of TET2 function is its most common co-mutated partner, it is logical to combine these mutations to establish how the mutations interact and collaborate to drive disease. Chen et al. (2015[225]) were first to create a cross using the Boston JAK2 V617F knock–in model[223] and the Moran-Crusio et al. TET2 knock-out model[135]. The resultant JAK2/TET2

double mutant mice presented similarly elevated haematocrit and platelets to the JAK2 V617F mice, alongside an increase in HSC numbers in the spleen and increased extramedullary haematopoiesis. The group showed, both *in vivo* and *in vitro*, that combining these two mutations improved the self-renewal of HSCs compared with JAK2 mutation alone.

Chen et al. performed gene expression profiling of bulk stem and progenitor cells (LSK) from WT, JAK2 V617F, TET2 knock-out and JAK2 V617F/TET2 knock-out mice by microarray. They reported that the individual genotypes had differing gene expression patterns, however it was concluded that the genes they identified were unlikely to drive the differences in function due to the lack of differentially expressed genes in multiple comparisons and minimal fold changes between mutant and WT. They attributed the main differences in gene expression to targets of STAT5A signalling (upregulated in JAK2 V617F and JAK2 V617F/ TET2 knock-out cells), a consequence of hyper-activation of the JAK/STAT signalling pathway. As the group performed gene profiling on bulk stem and progenitor cells they were unlikely to identify the drivers of this stem cell derived disease; less than 10% of the assayed LSK population are HSCs[138], and as JAK2 V617F has different effects on stem cells and progenitor cells (driving exhaustion of HSCs but expansion of progenitors[230]), it is understandable that the group were unable to identify the molecular drivers of the disease at the stem cell level.

Another JAK2 V617F TET2-mutant mouse cross has been performed by Kameda et al. (2015)[231]. This group crossed a JAK2 V617F transgenic model[232] (not covered in previous section because not a knock-in model) with a TET2 knock-down model[233] (20% TET2 mRNA expression compared to WT). From this cross they obtained JAK2 V617F TET2 knock-down E14.5 foetuses. FL cells from these animals were transplanted into lethally irradiated recipients. The recipients of double mutant cells exhibited an MPN-phenotype; splenomegaly, spleen and BM fibrosis, prolonged leucocytosis and shortened survival, a pMF-like disease.

The authors observed an increased percentage of LSK cells in the BM of TET2 knock-down and double mutant recipients, and an increased serial re-plating capacity of BM cells from the same mice. In competitive secondary transplantation (primary FL, secondary BM), double mutant cells gave rise to similar levels of chimerism to WT at

16 weeks post-transplantation, while in recipients of JAK2 V617F alone chimerism dropped to 5% (from 80% in primary) (at this point JAK2 recipients exhibited no MPN phenotype). The authors thus concluded that TET2 knock-down restores the JAK2 V617F self-renewal defect, conferring a competitive self-renewal advantage to double mutant cells and allowing an MPN phenotype to be sustained long-term.

Like Chen at el.[225], Kameda et al.[231] went on to do bulk gene expression of BM LSKs via microarray (BM from primary transplant). In JAK2 V617F and double mutant cells Kameda et al. also saw an enrichment in STAT5 target gene expression, alongside an enrichment in pre-erythroid forming colony gene signature. For the same reasons as Chen et al. this group were not able to detect the specific self-renewal genes altered in the JAK2, TET2 and double mutant HSCs, although they did report a reduction in HSC gene signature in JAK2 mutant cells – though this could be due to the reduced functionality of JAK HSCs, reduced number of JAK2 HSCs or reduced proportion of HSCs in LSK faction (due to increased progenitors).

TET2 is not the only gene that is commonly co-mutated with JAK2 V617F, mutations in other epigenetic regulators such as EZH2 have also been found to collaborate with JAK2 mutation in MPNs. A JAK2 V617F/EZH2 null double mutant mouse model has been developed,[234] this mouse is reported to have an MF like phenotype (with fibrosis of BM and spleen). Unlike JAK2/TET2 double mutant mouse models, which have haematocrit (HCT) and haemoglobin (HGB) elevated to the same extent as JAK2 mutant alone, JAK2/EZH2 double mutant animals have reduced HCT and RBC count compared to JAK alone, causing authors to draw the conclusion that loss of *Ezh2* impairs/inhibits erythropoiesis (JAK2/EZH2 also have reduced spleen sized when compared to JAK2 alone). However, like TET2, in the transplantation setting, loss of EZH2 enhances the repopulation capacity of JAK2 V617F HSCS.

Likewise, a JAK2 V617F/IDH2 R140Q double mutant mouse has been generated.[235] This mouse has elevated haematocrit, leukocyte count and splenomegaly. Competitive transplantation of BM from this double mutant mouse showed that, like loss of TET2, IDH2 R140Q leads to higher chimerism when combined with JAK2 V617F than JAK2 mutation alone. Together these double mutant mouse models show that mutations in

epigenetic regulators combine with JAK2 V617F to enhance the repopulative capacity of HSCs.

1.12 Aims of project

Malignant HSCs share many of the same cellular processes and pathways as normal HSCs, but some processes must be hijacked to enable increased proliferation and/or a differentiation block. Consequently, it becomes important to not think of malignant HSCs as a completely separate cell state to normal HSCs, but rather a very close molecular relative. In recent years a number of technical advances have allowed the development of single cell approaches, permitting subtle changes in individual characteristics to be assessed on a cell by cell basis. This has allowed the study of the mechanism by which mutations initiate and sustain disease at previously unobtainable resolution.

Since haematological malignancies often have more than one driver,[236] and several aspects of HSCs can be affected by each mutation (e.g., JAK2 mutations alter cell cycle status, proliferation, differentiation and HSC self-renewal), it remains difficult to account for the cause of individual properties of disease when studying single mutations in isolation. Consequently, numerous groups have developed model models for common combinations of mutations.[122,237,238] Studying combinations of mutations allows a full picture to be built up as to how each mutation influences disease characteristics. In studying the molecular effects of specific mutations on HSCs, mouse models are particularly useful, largely because it is possible to isolate murine HSCs at a much higher frequency compared to human, and additionally mouse models can be 100% mutant, thereby avoiding complications of intra-patient HSC heterogeneity in mutational state.

The aim of this thesis is to use both cellular and molecular single cell techniques to explore the functional and molecular consequences of the most common pair of mutations in MPN, JAK2 V617F and a loss of TET2, on murine HSCs, focusing on the impact of JAK2 V617F on HSC self-renewal.

2

Materials and Methods

2.1 Mice

JAK2 V617F knock-in mice[228] were crossed with Tet2 knock-out mice from Ko et al., (2011).[210] HSCs from single mutant mice (for brevity homozygous JAK2 V617F termed JAK HOM and homozygous TET2 null termed TET HOM) and double mutant (JAK HOM TET HET and JAK HOM TET HOM) animals were compared to wild-type littermate (for JAK HOM) or age-matched (for TET HOM and double mutant) controls. CALR mutant mice[201] were donated by AR Green, loss of CBP mice[239] were donated by BJP Huntly, NPM1 mutant mice[240] were donated by GS Vassiliou, and P53 knock-out mice[241] were donated by KJ Patel. c-Kit mutant C57BL/6$^{W41/W41\text{-}Ly5.1}$ (W41) or C57BL/6 mice were used as recipients for transplantation experiments. All recipient mice for transplantation were CD45.1, all donors CD45.2, and competitors were CD45.1/2 C57Bl/6 mice. All mice were bred and maintained at the University of Cambridge in microisolator cages and provided continuously with sterile food, water, and bedding. All mice were kept in specified pathogen-free conditions, and all procedures performed according to the United Kingdom Home Office regulations.

2.2 Isolation of ESLAM HSCs and SLAM HSPCs

Suspensions of BM cells were isolated from the femurs, tibiae, and iliac crests of mice. Red blood cell lysis was performed by treatment with ammonium chloride (STEMCELL Technologies, Vancouver, Canada (STEMCELL)), depletion of mature lineage cells was performed using EasySep mouse haematopoietic progenitor cell

enrichment kit (STEMCELL). ESLAM cells were isolated as described previously[24] using CD45 FITC (Fluorescein isothiocyanate) (Clone 30-F1,1 BD Biosciences, San Jose, CA, USA (BD Biosciences)), EPCR (CD201) PE (Phycoerythrin) (Clone RMEPCR1560, STEMCELL), CD150 PE-cy7 (Clone TC15-12F12.2, Biolegend, San Diego, USA (Biolegend)), CD48 APC (Allophycocyanin) (Clone HM48-1, Biolegend), and 7-Aminoactinomycin D (7AAD) (Thermo Fisher Scientific, Waltham, MA, USA (Thermo Fisher)). Although not used in the gating strategy, Sca1 (Ly-6A/E) Brilliant Violet (BV) 421 or Pacific Blue (PacB)(Clone D7, both Biolegend), and c-Kit (CD117) APC-cy7 (Clone 2B8, Biolegend) were included in some cases for BM analysis. Lineage markers (Lin) were, in some cases, stained for using the biotin conjugated EasySep Mouse Hematopoietic Progenitor Isolation Cocktail from the EasySep mouse haematopoietic progenitor cell enrichment kit (STEMCELL) and a secondary streptavidin-BV510 or BV605 (both Biolegend).

Example gating for ESLAM cells ($CD48^{-}CD150^{+}CD45^{+}EPCR^{+}$) can be seen in Figure 6 (results section 3.1), a similar strategy was used for HSPC (SLAM) isolation, but both EPCR positive and negative populations were sorted ($Lin^{-}CD48^{-}CD150^{+}CD45^{+}$). The cells were sorted on an Influx (BD Biosciences) using the following lasers; 405, 488, 561, and 640, and filter sets; 530/40 (for FITC), 585/29 (for PE), 750LP (for PE-Cy7 and APC-cy7), 670/30 (for APC and 7AAD), 460/50 (for PacB and BV421), 520/35 (for BV510), 610/20 (for BV605). When single ESLAM HSCs were required, the single-cell deposition unit of the sorter was used to place 1 cell into the wells of 96 well round bottom plates, each well having been preloaded with 50µL medium (described below).

2.3 *In vitro* cultures and clone size calculations

ESLAM HSCs were sorted and cultured in StemSpan (STEMCELL) containing foetal bovine serum (FBS; STEMCELL or Sigma-Aldrich, St. Louis, MO, USA (Sigma-Aldrich)), 300ng/mL stem cell factor (SCF; STEMCELL or Bio-techne, Abingdon, UK (Bio-techne)), 20 ng/mL interleukin (IL)-11 (STEMCELL or Bio-techne), 100 units/mL Penicillin and 100µg/mL Streptomycin (Sigma-Aldrich), 2mM L-Glutamine (Sigma-Aldrich), and 0.1mM 2-Mercaptoethanol (Thermo Fisher) (together HSC medium). Cells were cultured at 37°C, 5% CO2. Cell counts were performed daily, and

cell cycle kinetics determined for the first and second division by visual inspection, manually scoring each well as having 1, 2, or 3-4 cells. At 10 days, clones were estimated to be very small (less than 50 cells), small (50–500 cells), medium (500–10,000 cells), or large (10,000 or more cells). Clone size estimates were previously validated using fluorescent counting beads.[230] Ten-day clones were stained with Sca1 (Ly-6A/E) PacB (Clone E13-161.7, Biolegend), Mac1 (CD11b) BV605 (Clone M1/70, Biolegend), Gr1 (Ly-6G/Ly-6C) PE-cy7 (Clone RB6-8C5, Biolegend), c-Kit (CD117) APC-cy7 (Clone 2B8, Biolegend), CD45.2 FITC (Clone 104, Biolegend), and 7-AAD (Thermo Fisher), example gating in results (Figure 6). To enumerate cells, a defined number of fluorescent beads (Trucount Control Beads, BD Biosciences) were added to each well and each sample was back-calculated to the proportion of the total that were run through the cytometer. Small clones were not assessed individually by flow cytometry due to low numbers of cells. Flow cytometry was performed on an LSRFortessa (BD Biosciences) and all data were analysed using FlowJo (Treestar, Ashland, OR, USA).

2.4 Isolation of patient stem and progenitor cells

Fresh venous blood samples (40-60mL collected in S-Monovette Lithium-Heparin tubes (Sarstedt, Nümbrecht, Germany (Sarstedt)) were collected from MPN patients with JAK2 and TET2 mutations, diagnosed according to British Committee for Standards in Haematology (BCSH) guidelines. The study was approved by the Cambridge and Eastern Region Ethics Committee, and was carried out in accordance with the principles of the Declaration of Helsinki. Informed written consent was obtained from all patients before participating. Mononuclear cells (MNCs) were isolated using Lymphoprep (STEMCELL) according to the manufacturer's instructions. MNCs were washed with and stored at 4°C overnight in phosphate-buffered saline (PBS, Sigma-Aldrich) supplemented with 10% FBS (STEMCELL or Sigma-Aldrich). In the morning, cells were washed with PBS (Sigma-Aldrich). For those that were grown *in vitro* and for frozen MNCs, cells were spun at 300xg for 7 minutes and the resulting pellets were treated with 50µL of 1 mg/mL DNaseI (STEMCELL) to avoid clumping. Cells were then depleted of differentiated haematopoietic cells using the EasySep human haematopoietic progenitor enrichment kit (STEMCELL) with the following modifications: All samples were processed in 500µL of recommended

medium, and the sample was only taken through one round of separation. Cells were then stained with antibodies to isolate the various progenitor compartments: CD34 Per-CpCy5.5 or PE-cy7 (clone 581, both Biolegend), CD38 FITC (clone HIT2, BD Biosciences), CD90 APC (clone 5E10, BD Biosciences) and CD45RA Horizon V450 (clone HI100, BD Biosciences).

$CD34^+CD38^-$ cells were sorted into microcentrifuge tubes containing 500µL of PBS/10%FBS. HSCs were isolated as $Lin^-CD34^+CD38^-CD45RA^-CD90^+$ using a BD Biosciences Influx sorter equipped with 355nm, 405nm, 488nm, 561nm, and 640nm lasers sorted directly into individual wells of a 96-well round bottomed plate. HSCs were sorted as single cells into StemSpan medium (STEMCELL) supplemented with the cc100 cytokine cocktail (STEMCELL), 100 units/mL Penicillin and 100µg/mL Streptomycin (Sigma-Aldrich), 2mM L-Glutamine (Sigma-Aldrich), and 0.1mM 2-Mercaptoethanol (Thermo Fisher). For proliferation studies HSC derived clones were counted visually, on day 10 clones were classified by diameter as very small (<50 cells), small (50–500 cells), medium (500–10,000 cells), or large (10,000 or more cells). To improve genotyping accuracy, single HSC-derived cultures were harvesting and plated into a secondary CFC assay using Methocult 04435 (STEMCELL), the resultant colonies were genotyped for TET2 and JAK2 mutation status.

2.5 Bone Marrow Transplantation Assays

Donor cells were obtained from mice of various genotypes, all donor mice were on a C57BL/6J background (CD45.2). For whole BM non-competitive transplantations $5x10^5$ donor cells were transplanted alone. For limiting dilution transplantations 100,000, 50,000 or 10,000 WT or JAK HOM TET HOM BM cells were transplanted alongside 100,000 competitor cells from WT C57BL/6J (CD45.1/CD45.2) mice. For whole BM competitive transplantations $5x10^5$ donor cells were transplanted alongside $5x10^5$ competitor cells obtained from WT C57BL/6J (CD45.1/CD45.2) mice.

For whole BM transplantation assays, recipients were adult W41 mice as described previously.[242,243] Mice were irradiated with a single dose (400 cGy) by Caesium irradiation and all transplants were performed by intravenous tail vein injection using a 29.5G insulin syringe. For competitive transplantation, on average, host cell recovery accounted for 30% of the total cells (15-60%) but since a known number of

genotypically distinct competitor cells were transplanted, host numbers were not required for calculations of relative chimerism between donor and competitor cells (donor/(donor+competitor)). For secondary transplantation, whole BM was obtained from primary recipients and $1x10^7$ cells, containing a mixture of recipient, (competitor,) and donor-derived cells, were transplanted.

2.6 Peripheral Blood and Bone Marrow Analysis

For all transplantation assays, serial bleeds were performed, for these peripheral blood samples were collected in EDTA (Ethylenediaminetetraacetic acid) coated microvette tubes (Sarstedt) from the tail vein of mice at regular intervals following transplantation. Peripheral blood cell counts were performed using a Woodley ABC blood counter (Woodley Equipment, Bolton, UK). Red cell lysis was performed using ammonium chloride (STEMCELL) and samples were subsequently analysed for repopulation levels by flow cytometry, as described previously.[27,242] For flow cytometry the following antibodies were used; Gr1 (Ly-6g/Ly6c) PacB or BV421 (Clone RB6-8C5, both Biolegend), Mac1 (CD11b) FITC or BV605 (Clone M1/70, both Biolegend), CD3 PE (Clone 17A2, Biolegend), B220 (CD45R) APC (Clone RS3-6B2, Biolegend), CD45.1 AlexaFluor700 (AF700) (Clone A20, Biolegend), and CD45.2 APC-cy7 or FITC (Clone 104, both Biolegend). 7-AAD (Thermo Fisher) was used as a live/dead stain. Example lineage gating Figure 5.

For transplants for which secondary transplantation was performed (competitive and non-competitive whole BM transplantation), at the end of the primary and secondary

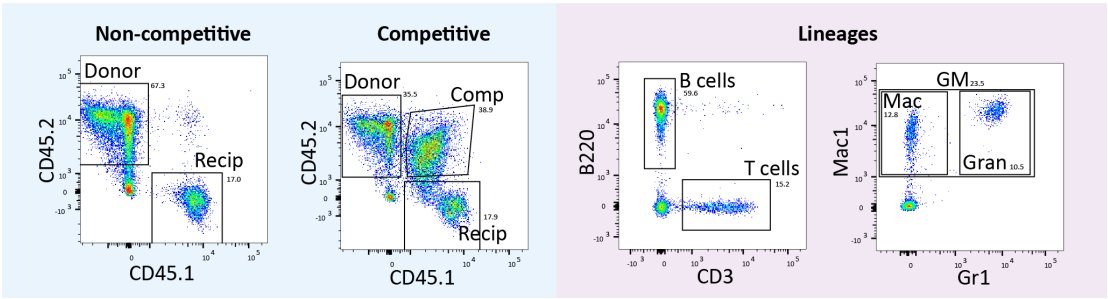


Figure 5 Example gating for post-transplantation peripheral blood
Example of donor population gating from competitive and non-competitive transplantation (left), and donor lineages (right). B- and T-cells, granulocytes and macrophages are gated from live donor cells ($7AAD^-$, $CD45.2^+$, $CD45.1^-$). Cells which did not fall into any of these lineage gates were termed 'other'.

transplantation time course, mice were sacrificed and BM harvested. BM was analysed by flow cytometry for lineage contribution (using the same panel as peripheral blood) and stem and progenitor cell contribution. For assessment of stem and progenitor cell contribution the following antibodies were used; CD45.1 AF700 (Clone A20, Biolegend), CD45.2 APC-cy7 or FITC (Clone 104, both Biolegend), EPCR PE (Clone RMEPCR1560, STEMCELL), CD150 PE-cy7 (Clone TC15-12F12.2, Biolegend), CD48 APC (Clone HM48-1, Biolegend), Sca1 (Ly-6A/E) BV605 (Clone D7, Biolegend), c-Kit (CD117) APC-cy7 (Clone 2B8, Biolegend). As before lineage markers were stained for using the biotin conjugated EasySep Mouse Hematopoietic Progenitor Isolation Cocktail from the EasySep mouse haematopoietic progenitor cell enrichment kit (STEMCELL) and a secondary streptavidin-BV510 (Biolegend), and 7-AAD (Thermo Fisher) was used as a live/dead stain.

2.7 Gene-expression profiling

Single-cell gene expression analysis of HSC associated genes (Table 1) was performed as described previously.[244,245] The 48 TaqMan assays (Thermo Fisher, primer IDs listed in Table 3) were pooled to a final concentration of 0.2X (a 1:100 dilution of each assay in TE buffer (Thermo Fisher)). Single HSCs were sorted directly into 96-well plates containing 5µL CellsDirect 2X reaction mix (Thermo Fisher), 0.1µL SUPERase RNase inhibitor (Thermo Fisher), 2.5µL 0.2X assay mix, 1.2µL TE buffer (Thermo Fisher) and 1.2µL SuperscriptIII/Platinum Taq (Thermo Fisher). Cell lysis and sequence-specific reverse transcription were performed at 50°C for 15 minutes. The reverse transcriptase was inactivated by heating to 95°C for 2 minutes. Subsequently, in the same wells, cDNA went through sequence-specific amplification by denaturing at 95°C for 15s and annealing and amplification at 60°C for 4 min for 22 cycles, followed by a final 60°C for 4 min, then stored at -80°C. These preamplified products were diluted 1:5 with TE buffer prior to analysis on the BioMark HD (Fluidigm, San Francisco, CA, USA (Fluidigm)). 48.48 dynamic arrays were run; each with 6 no template controls, 1 positive control well (containing 10 cells) and 41 single cells. 48.48 dynamic arrays chips were loaded as per the manufacturer's instructions using an IFC Controller MX (Fluidigm), and run on a BioMark HD (Fluidigm) using the GE 48.48 standard v1 protocol (95°C for 10 minutes; 40 cycles of 95°C for 15 s and 60°C for 60 s).

Table 3 List of genes for which expression was analysed by single cell multiplex qPCR

Gene	TaqMan order ID
Bmi1	Mm00776122_gH
Bptf	Mm01251151_m1
Cbfa2t3h	Mm00486780_m1
Cbx7	Mm00520006_m1
Cdkn2a	Mm00494449_m1
Csf1r	Mm01266652_m1
Dnmt3a	Mm00432881_m1
Egfl7	Mm00618004_m1
Eif2b1 (housekeeping)	Mm01199614_m1
Erg	Mm01214246_m1
Ezh2	Mm00468464_m1
Fli-1	Mm00484409_m1
Foxo3a	Mm01185722_m1
Gata1	Mm00484678_m1
Gata2	Mm00492300_m1
Gata3	Mm00484683_m1
Gfi1	Mm00515855_m1
Gfi1b	Mm00492318_m1
Hhex	Mm00433954_m1
Hoxa5	Mm00439362_m1
Hoxa9	Mm00439364_m1
Hoxb4	Mm00657964_m1
Ikzf1	Mm01187882_m1
Itga2b	Mm00439768_m1
Kit	Mm00445212_m1
Lmo2	Mm01281680_m1
Lyl1	Mm01247198_m1
Mecom	Mm01289155_m1
Meis1	Mm00487659_m1
Mitf	Mm01182480_m1
Mpl	Mm00440310_m1
Myb	Mm00501741_m1
Nfe2	Mm00801891_m1
Pbx1	Mm04207617_m1
Polr2a (housekeeping)	Mm00839493_m1
Prdm16	Mm00712556_m1
Procr	Mm00440993_mH
Runx1	Mm01213405_m1
Sfpi1	Mm00488142_m1
Sh2b3	Mm00493156_m1
Smarcc1	Mm00486224_m1
Smarcc2	Mm01159912_m1
Tal1	Mm01187033_m1
Tcf7	Mm00493445_m1
Tet2	Mm00524395_m1
Trib3	Mm00454879_m1
Ubc (housekeeping)	Mm01201237_m1
Vwf	Mm00550376_m1

Ct values were obtained from the system's software (BioMark Real-time PCR Analysis; Fluidigm). ΔCt values were calculated as previously described[245] by cell-wise normalisation to the mean expression level of two housekeeping genes (Ubc and Polr2a). All housekeepers, Cdkn2a, Eif2b1, Meis1, Tet2, and Egfl7 were removed from the dataset for downstream analysis. Cdkn2a was not expressed in any of the cell types, Egfl7, Mitf, Cbx7, and Eif2b1 did not pass quality control in any genotype, Meis1 was excluded in the TET2 single and double-mutant genotypes due to aberrant signal in no template controls, and Tet2 was excluded since the TET2 knock-out cells had no expression and clustering would have been unfairly biased. Hierarchical clustering was performed in R (http://www.r-project.org) using the hclust package and heatmap.2 from the gplots package using Spearman rank correlations and ward linkage. T-distributed stochastic neighbour embedding (t-SNE) was performed in R (http://www.rproject.org) using the Rtsne package. Principal component analysis (PCA) was performed using the prcomp function.

2.8 Lentivirus production

Lentivirus production was performed with HEK Lenti-X 293T cells (Takara Bio Inc, Kusatsu, Japan (Takara)) using TransIT-LT1 (Mirus Bio, Madison, WI, USA). Cells were co-transfected with the plasmid of interest in a pCCL-c-MNDUS-X2-PGK-EGFP backbone, pMDLg/pRRE (Gag-Pol), pRSV-Rev (Rev), pMD2.G (VSV-G envelope) pAdVAntage (Promega, Madison, WI, USA; increases lentirvirus production[246]) (all kindly provided by Priyanka Tibarewal, University of Cambridge).

HEK Lenti-X 293T cells were cultured in low glucose Dulbecco's modified Eagle's medium with GlutaMAX and sodium pyruvate (Thermo Fisher), supplemented with 10% FBS (Sigma-Aldrich), 2mM L-Glutamine (S Sigma-Aldrich), 100 units/mL penicillin and 100ug/mL streptomycin/mL (Sigma-Aldrich). 16 hours after liposomal transfection, cell culture medium was supplemented with 10-12.5 mM sodium butyrate (Sigma-Aldrich). After 6 hours, cells were washed with PBS (Sigma-Aldrich) and cell culture medium replaced (Iscove's Modified Dulbecco's Medium (IMDM) with GlutaMAX, sodium pyruvate and 25mM HEPES buffer (Thermo Fisher), supplemented with 10% FBS (Sigma-Aldrich), 2mM L-Glutamine (Sigma-Aldrich), 100 units/mL penicillin and 100ug/mL streptomycin/mL (Sigma-Aldrich)). 22 hours

later, the virus-containing supernatant was collected and filtered through 0.45µm syringe filters. Virus was concentrated by ultracentrifugation (using a Beckman Coulter Optima XPN 80K, Beckman Coulter, Brea, CA, USA) at 25,000rpm for 2 hours (with breaks off), before being resuspended in IMDM (w 25mM HEPES, STEMCELL) and stored at -80ºC.

2.9 Overexpression assays

Small pools (1500-3600 cells) of $CD45^+Lin^-CD150^+CD48^-$ haematopoietic stem and progenitor cells (HSPCs) were isolated (as described above) and split between wells of a 96 well plate coated with RetroNectin (Takara). Following their isolation, cells were kept in 50µL of HSC medium (described above) and were supplemented with polybrene (Sigma-Aldrich) and a lentivirus equipped with the gene of interest in a pCCL-c-MNDUS-X2-PGK-EGFP backbone. Plates were centrifuged at 600 x g for 30 minutes, 30ºC, to promote infection, before being transferred to a 37ºC incubator. The following day, a further 150µL of HSC medium was added to dilute the polybrene and virus. Three days after infection, cells were resorted for green fluorescent protein (GFP) expression and viability (i.e. GFP^+ $7AAD^-$). Cells (300-2000) were transplanted into sublethally irradiated (400cGy) CD45.1 W41 or lethally irradiated (1100cGy) CD45.1 C57Bl6/J recipient mice and monitored for donor chimerism and disease phenotype as described above. At the end of the transplantation time course recipient mice were sacrificed and their spleens measured and weighed to assess for splenomegaly.

2.10 Primary patient sample and mouse post-transplantation gene expression

RNA was extracted from sorted human $CD34^+CD38^-$ cells, and mouse SLAM cells ($CD45^+CD48^-CD150^+$) using the PicoPure RNA isolation kit (Thermo Fisher) as per the manufacturer's instructions, with on column DNase treatment using RNase-free DNase set (Qiagen, Hilden, Germany), and elution in 11µL. First strand cDNA synthesis was performed using SuperScript III First-Strand Synthesis System for RT-PCR (Thermo Fisher), using Random Hexamers, as per the manufacturer's guidelines; 10µL of template cDNA was used, with 1.25µL of Random hexamers and 1.25µL dNTP mix (as 10µL of cDNA was used, the volumes of components were scaled to 1.25x the recommend volumes). qPCR was carried out using TaqMan probes (probe IDs for

mouse as listed in Table 3, for human listed in Table 4) and TaqMan universal master mix, as per the manufactures instructions, in a 20µL reaction with 2µL of cDNA. qPCR assays were run and analysed using the Applied Biosystems ViiA 7 system (Thermo Fisher). The expression of individual genes relative to a housekeeping gene (GAPDH) was calculated as $R = 2^{-[CT\ target - CT\ housekeeping]}$.[247]

Gene	TaqMan order ID
Meis1	Hs00180020_m1
Bmi1	Hs00180411_m1
Runx1	Hs02558380_s1
Pbx1	Hs00231228_m1
GAPDH	Hs02786624_g1

Table 4 List of genes for which expression was analysed in human samples by qPCR

2.11 Statistical Analyses

For comparison of cell division kinetics and colony size two-tailed Fishers Exact test was used, for calculating cluster enrichment scores Chi-squared test was performed (two-tailed, yates correction), for both the web based calculator at https://www.graphpad.com/quickcalcs/contingency1/ was used. For all other p values reported, an unpaired Student's t-test (Microsoft Excel) was used. For comparison of blood parameters t-tests were one-tailed with unequal distribution. For all others two-tailed with unequal distribution.

Stem cell frequency from limiting dilution transplantation was calculating using the webtool at http://bioinf.wehi.edu.au/software/elda/.

Results

3

MPN HSCs *in vitro*

I began this chapter by focusing on the effect of JAK2 and TET2 mutations, both individually and in combination, on murine HSCs in culture. I then explored if the phenotypes displayed by the JAK2 and TET2 mutant HSCs are shared with other common mutations in myeloid malignancy with similar characteristics. Finally, I validated the effects of JAK2 and TET2 HSC *in vitro* using data from the culture of human MPN patient stem and progenitor cells. The human data in section 3.3 was from work carried out by Christina Ortmann and David Kent [190]

The JAK2 mutant and TET2 mutant models used throughout this thesis are a homozygous human JAK2 V617F under the control of the mouse endogenous promoter,[228] and a partial or complete TET2 knock-out.[210] For the sake of brevity, homozygous JAK2 V617F animals will be termed JAK HOM, homozygous loss of TET2 animals TET HOM, heterozygous loss of TET2 TET HET, and double mutant animals the appropriate combination of these terms (e.g. JAK HOM TET HOM).

3.1 JAK2 V617F drives increased proliferation and differentiation of HSCs *in vitro*, irrespective of TET2 status

The study of HSCs often begins with the isolation of purified populations of HSCs. Throughout this thesis $CD45^{+}EPCR^{+}CD48^{-}CD150^{+}$ (ESLAM) sorting strategy (Figure 6) was used to isolate enriched population of HSCs, this population has been shown to be 40%-50% long-term repopulating HSCs by single cell transplantation studies.[24]

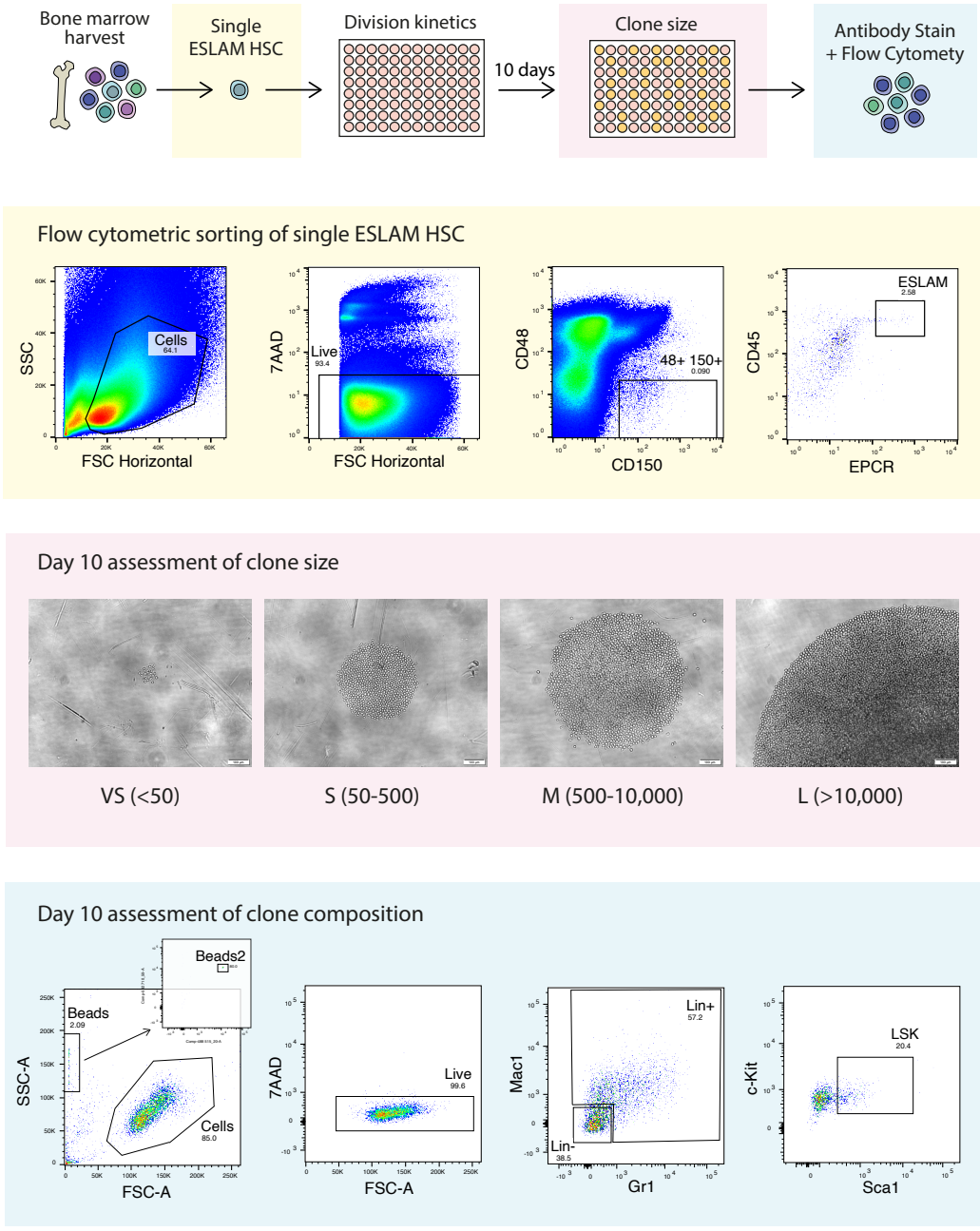


Figure 6 Schematic of single cell in vitro culture and representative images
Single CD45$^+$EPCR$^+$CD48$^-$CD150$^+$ (ESLAM) cells (representative gating, yellow box) were sorted into individual wells of 96 well plate, cultured for 10 days in StemSpan with 10% FCS, 300ng/mL SCF, and 20 ng/mL IL-11, and assessed for proliferation, cell cycle kinetics, and differentiation. Representative images very small (VS), small (S), medium (M) and large (L) clones at day 10; pink box. Representative gating of a day 10 clone (blue box); gating for live cells (7AAD$^-$), differentiated cells (Lin$^+$; Mac1$^+$ and/or Gr1$^+$), and stem and progenitor cells (LSK; Lin$^-$, c-Kit$^+$, Sca1$^+$).

Hyperproliferation is a key trait of leukaemic clones, and as discussed in the introduction (section 1.6.1) there are a number of assays that can be used to assess hyperproliferation *in vitro*. The single cell liquid culture system was used here; this assay allows division kinetics from a single starting HSC to be assessed by daily counts, the clone size and lineage composition to be assessed by flow cytometry, and permits these functional findings to be coupled with initial HSC cell surface marker expression.

To determine the effect of JAK2 V617F expression and TET2 knock-out on HSCs in culture, single ESLAM HSCs were cultured in liquid culture conditions that support HSC self-renewal; briefly StemSpan medium supplemented with FBS, L-glutamine, penicillin, streptomycin, 300ng/mL SCF, and 20 ng/mL IL-11,[248] and daily counts performed. After 9-10 days in culture, colony size was assessed visually (representative images of clone sizes in Figure 6); clones with fewer than 50 cells were grouped as very small clones, clones with between 50 and 500 cells were classified as small clones, clones containing between 500 and 10,000 cells were termed medium clones, and clones with over 10,000 cells dubbed large clones. The composition of each individual clone was assessed by flow cytometry (representative flow cytometry plots, Figure 6), and absolute cell numbers determined using Trucount beads (BD Bioscience).

Comparison of the cell division kinetics between WT and JAK HOM ESLAM HSCs showed JAK HOM HSCs exited quiescence more quickly (Figure 7). JAK HOM HSCs having a shorter time to first cell division; after 2 days in culture 87.67% JAK HOM cells had divided vs 60.31% WT cells (Figure 7 top left, $p<0.0001$). This was followed by a shorter time to subsequent cell division, with an increased proportion of JAK HOM HSCs (45.5%) having completed their second division after 2 days in culture compared with WT controls (16.6%) (Figure 7 top right, $p < .0001$). Assessment of clonal progeny of individual HSCs showed that colonies from JAK HOM HSCs were on average larger (Figure 7 bottom left, p VS $p<0.0001$, S $p<0.0001$, M $p=0.0003$, L $p<0.0001$) and more differentiated (Figure 7 bottom middle, $p<0.0001$) and contained proportionally fewer progenitor cells (LSK; Figure 7 bottom right, $p<0.0001$) than clones grown from HSCs obtained from WT littermate controls.

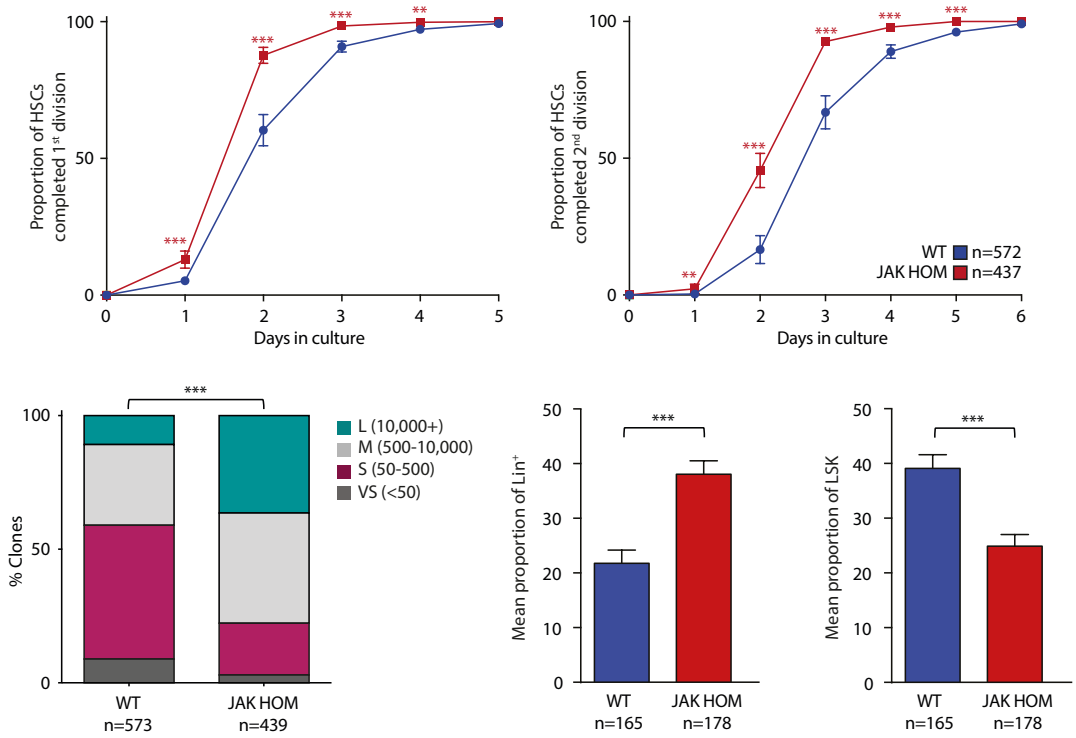


Figure 7 JAK2 V617F alone drives increased proliferation and differentiation of HSCs in vitro

Daily cell counts reveal that JAK HOM HSCs (red line, N=8) have accelerated cell cycling when compared with WT HSCs (blue line, N=8) as indicated by a shorter time to first and second division, $p<0.0001$ at 2 days for both. JAK HOM HSCs give rise to on average larger clones than WT littermates, $p<0.0001$. JAK HOM HSCs (red bars) give rise to an increased number of differentiated cells (Lin^+)($p<0.0001$) and a reduced number of stem/progenitor cells (LSK)($p<0.0001$). Error bars are standard error of the mean (SEM).

When loss of TET2 alone was studied, no change was seen from WT controls in time to exit from quiescence (Figure 8 top left, at 2 days p=0.903) or time to subsequent cell division (Figure 8 top right, at 3 days p=0.381). Supporting the finding that cell division kinetics are unaffected by loss of TET2, there was also no difference between TET HOM and WT day 9 clone size (VS p=0.029, S p=0.953, M p=0.020, L p=0.559). Likewise, no change was seen in the frequency of lineage positive (p=0.369) or LSK (p=0.951) cells within the TET knock-out clones, relative to WT clones (Figure 8).

In order to determine how these two mutations affect HSC proliferation when combined - whether double mutant HSCs have a JAK V617F- or TET knock-out -like 'phenotype' or somewhere in between - HSCs from double mutant mice (homozygous for JAK2

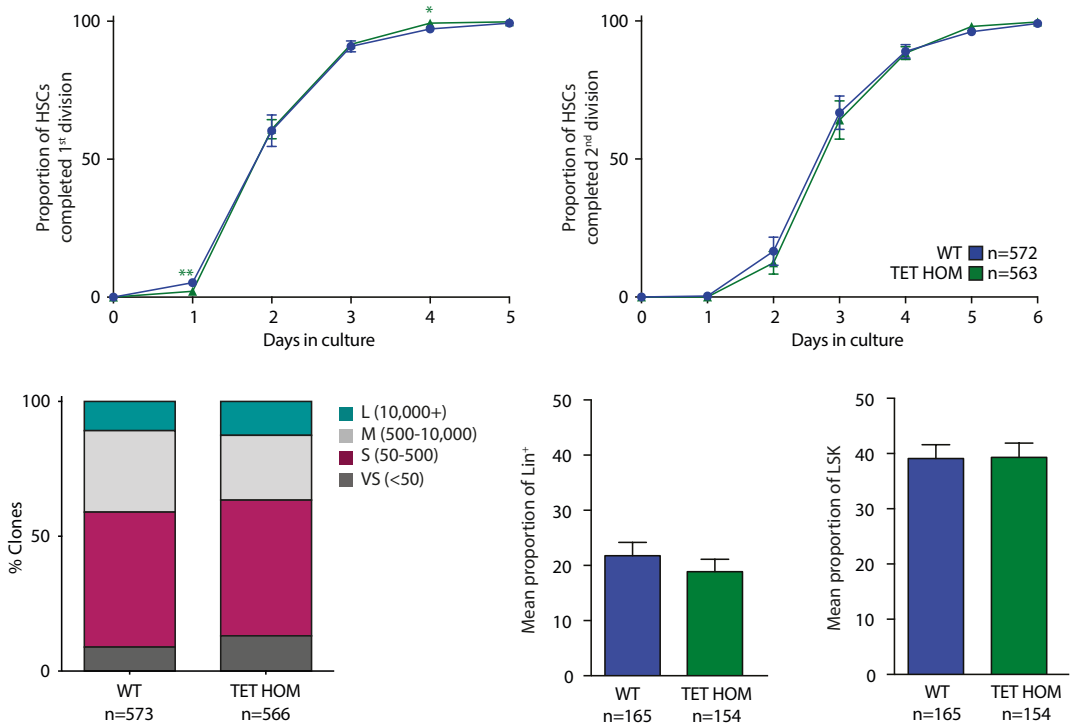


Figure 8 Loss of TET2 alone does not alter proliferation and differentiation of HSCs in vitro

TET HOM alone HSCs (green line, N=5) do not have altered cell cycling compared to WT HSCs (blue line, N=8), for 1st division at day 2 p=0.903, and for 2nd division at day 3 p=0.381. TET HOM HSCs do not give rise to more large clones than WT HSCs, p=0.559. TET2 clones (green bars) contain similar proportions of Lin+ (p=0.369) and LSK (p=0.951) cells to WT clones (blue bars). Error bars are SEM.

V617F and either heterozygous or homozygous for loss of TET2) were tested in the same assay.

When heterozygous loss of TET2 was combined with homozygous JAK2 V617F (JAK HOM TET HET), entry into cell cycle was accelerated (Figure 9 top left, at 2 days 89.7% of JAK HOM TET HET HSCs had divided vs 60% of WT $p > 0.0001$), cell cycle transit time was increased (Figure 9 top right, at day 3 $p>0.0001$), and average clone size was increased (Figure 9 bottom left, VS, S, L all $p<0.0001$, M $p=0.0003$), similar to the phenotypes observed in JAK2 V617F homozygous HSCs. Clones derived from JAK HOM TET HET HSCs also had an increased proportion of mature lineage marker-positive cells (Figure 9 bottom middle, $p=0.004$) and a decreased proportion of primitive progenitor (LSK) cells (Figure 9 bottom right, $p>0.0001$).

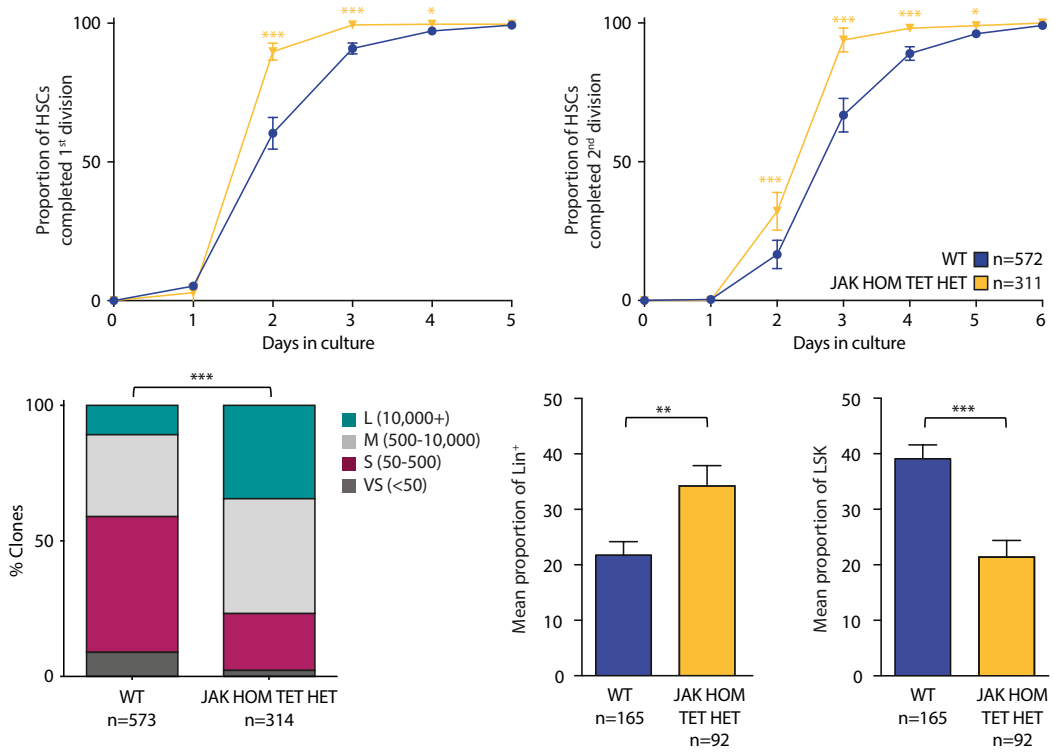


Figure 9 Homozygous JAK2 V617F combined with heterozygous loss of TET2 accelerates HSC proliferation in vitro

JAK HOM TET HET HSCs (yellow line, N=4) have accelerated cell division kinetics when compared with WT HSC (blue line N=8) as indicated by a shorter time to first ($p<0.0001$ day 2) and second division ($p<0.0001$ day 3). JAK HOM TET HET HSCs give rise to on average larger clones than WT HSCs, $p<0.0001$. JAK HOM TET HET HSCs (yellow bars) give rise to an increased number of differentiated cells (Lin^+, $p=0.004$) and a reduced number of stem/progenitor cells (LSK, $p>0.0001$), error bars are SEM.

Similarly, when complete loss of TET2 was combined with homozygous JAK2 V617F (JAK HOM TET HOM), time to first division was reduced (Figure 10 top left, at day 2 p>0.0001 at 2), as well as time to second division (Figure 10 top right, at day 3 p>0.0001). Just as in JAK HOM alone and JAK HOM TET HET, average clone size was increased (Figure 10 bottom left, VS p=0.0005, S p<0.0001, M p=0.078, L p<0.0001), the proportion of lineage positive cells was increase (Figure 10 bottom middle, p=0.018), and proportion of LSK cells decreased relative to WT clones (Figure 10 bottom right, p=0.001). From these models it can be concluded that compound loss of TET2 does not reverse the JAK2 V617F-induced cycling and hyperproliferation phenotypes observed in single *in vitro* HSC functional assays.

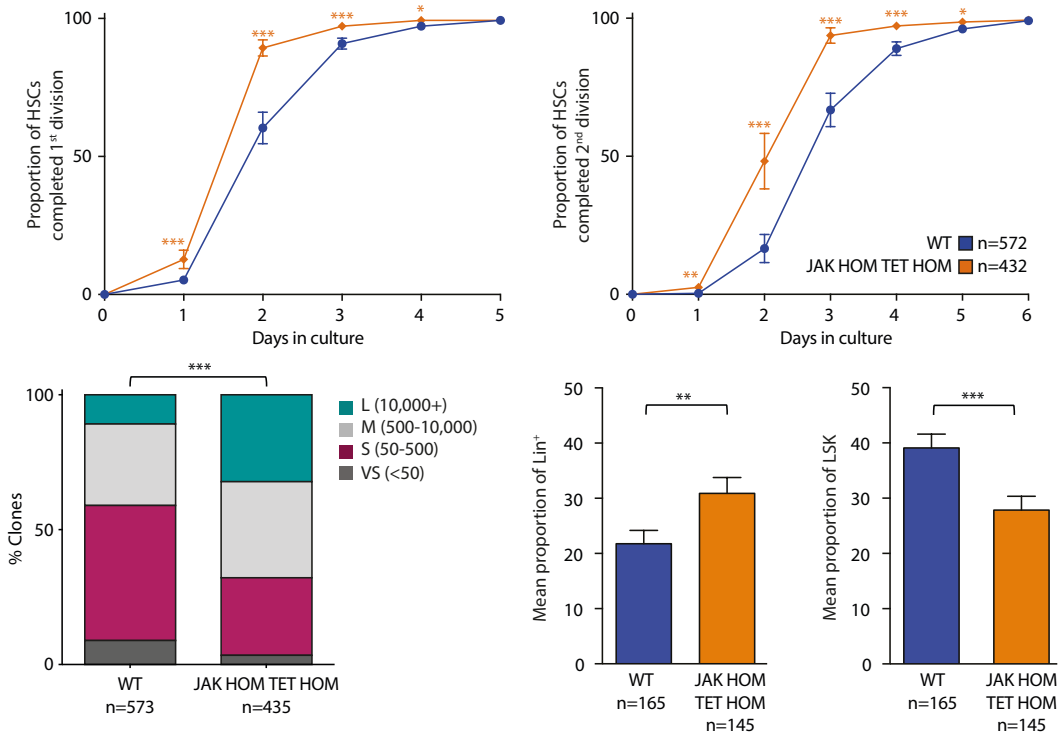


Figure 10 Homozygous JAK2 V617F combined with homozygous loss of TET2 accelerates HSC proliferation in vitro

JAK HOM TET HOM HSCs (orange line, N=6) have accelerated cell division kinetics when compared with WT HSCs (blue line, N=8) as indicated by a shorter time to first ($p<0.0001$ day 2) and second division ($p<0.0001$ day 3). JAK HOM TET HOM HSCs give rise to on average larger clones than WT HSCs, $p<0.0001$. JAK HOM TET HOM HSCs (orange bars) give rise to an increased number of differentiated cells (Lin^+, $p=0.018$) and a reduced number of stem/progenitor cells (LSK, $p=0.001$). Error bars are SEM.

3.2 Impact of myeloid malignancy mutations on murine HSC kinetics *in vitro*

To assess whether genes commonly found to be mutated in patients with myeloid malignancies cause generalised effects on HSCs, for example do all signalling mutations accelerate cell cycle? ESLAM HSCs from a number of models of myeloid malignancy were cultured *in vitro* and assessed for their cell cycle kinetics, and subsequent day 10 colony size.

3.2.1 CALR mutant HSCs behave similarly to JAK2 mutant HSCs *in vitro*

CALR is one of the three genes most commonly mutated in MPNs, alongside JAK2 and MPL. Mutations in all three genes lead to hyper-activation of the JAK/STAT

signalling pathway (see introduction section 1.10.1 signalling mutations in MPNs). As mentioned in the introduction, a CALR mutant mouse mimicking the CALR del52 mutation, commonly seen in patients, has been developed.[201] Both homozygous and heterozygous versions of this model have been reported to have an elevated number of ESLAM HSCs, but neither have a competitive advantage upon primary or secondary transplantation. To determine the similarity between CALR mutation and JAK2 mutation on the *in vitro* capabilities of HSCs, homozygous (CALR HOM) and heterozygous (CALR HET) CALR mutant ESLAM HSCs were cultured in the same manner as above (Figure 6).

When compared to WT cells, neither CALR HET nor HOM HSCs were significantly different in their time to first cell division (at 2 days HET p=0.510, HOM p=0.434, Figure 11 left). However, CALR HOM HSCs were substantially accelerated in their time to second cell division (98.88% CALR HOM divided vs 88.17% of WT at day 3, p=0.005, Figure 11 middle). CALR HET HSCs were not significant different to WT in their time to second division (at day 3 p=0.825). After 10 days in culture, compared to WT HSCs, CALR HET HSCs had given rise to a larger proportion of large colonies (36.36% CALR HET vs 18.28% WT, p=0.007), but the proportion of large colonies from CALR HOM HSC was higher still (66.29%, p < 0.0001) (Figure 11 right).

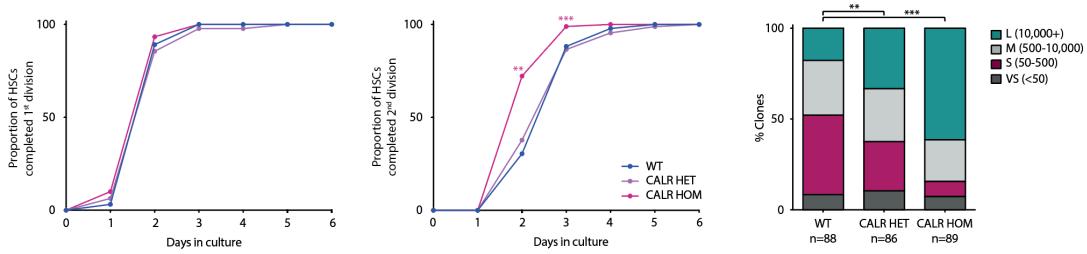


Figure 11 Homozygous and heterozygous CALR mutant HSCs give rise to on average larger clones after 10 days in culture

While neither CALR HET (purple line, N=1) or CALR HOM (pink line, N=1) have accelerated entry into cell cycle (first cell division, left, at 2 days HET p=0.510, HOM p=0.434,) compared with WT HSC (blue line, N=1). CALR HOM HSCs have a reduced time to second cell division (day 2 p<0.0001). Both CALR HET and HOM give rise to larger clones after 10 days in culture (HET p=0.007, HOM p < 0.0001).

3.2.2 CBP null HSCs have delayed cell division kinetics

Inactivating mutations in Cyclic AMP Response Element Binding (CREB)-binding protein (CBP, also known as CREBBP) have been reported in numerous haematological malignancies including ALL and B-cell lymphoma.[249,250] CBP is a protein of many functions and can influence gene expression in a number of ways[251]; firstly, in its role as a molecular scaffold, CBP facilitates the coupling of transcription factors with transcriptional machinery. Secondly, CBP remodels chromatin by histone acetylation, thereby opening chromatin and facilitating transcription. Thirdly, CBP can act to acetylate, and consequently activate, associated proteins including p53 (discussed below) and GATA-1, which have well established roles in haematopoietic differentiation.[251]

As inactivating CBP mutations are common in cancers, CBP knock-out mice have been developed. Due to CBP having an essential role in the development of HSCs, and deletion of CBP being embryonic lethal, conditional CBP knock-out mice have been developed to investigate the role of CBP in adult tissues. A conditional CBP model (conditional deletion of exon 9[252]) has been crossed with Mx1-Cre to generate a model with conditional ablation of CBP upon pIpC treatment.[239] This mouse has been shown to have a myeloid (Granulocyte/Macrophage; GM) differentiation bias and self-renewal defect, as determined by serial competitive transplantation.[239]

Unlike the previously described models which have equal or accelerated cell division kinetics compared to WT, HSCs from the CBP mutant model have slower entry into

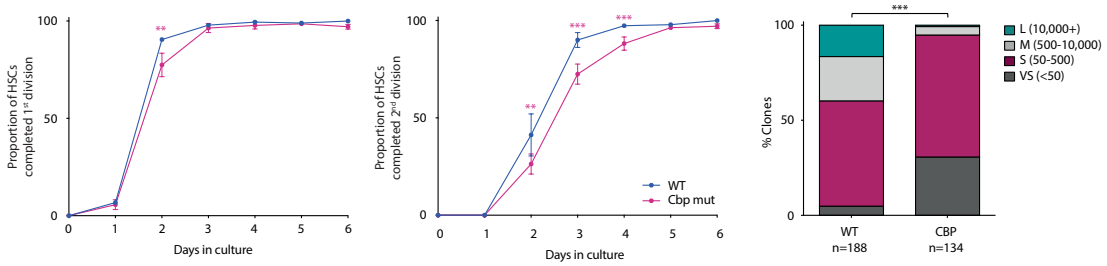


Figure 12 CBP mutation delays HSCs proliferation in vitro

CBP mutant HSCs (pink line, N=2) had a delayed entry into cell cycle (time to first division) compared with WT HSCs (blue line, N=2), (at 2 days p=0.002), and a prolonged time to subsequent cell division (at 3 days p=0.0001). After 10 days in culture CBP mutant clones were on average smaller than equivalent WT clones (VS, M, L p=0.0001, S p=0.217). Error bars are SEM.

cell cycle compared to WT (Figure 12 left, at 2 days 77.44% of CBP HSCs had divided vs 90.48 % of WT HSCs, p=0.002) and delayed time to subsequent cell divisions (Figure 12 middle, at 3 days 72.46% of CBP HSCs had completed 2 divisions vs 90 % WT, p=0.0001), resulting in smaller average clone sizes after 10 days in culture (Figure 12 right, VS, M, L p=0.0001, S p=0.217).

3.2.3 NPM1 mutant HSCs have delayed cell division kinetics

Nucleophosmin 1 (NPM1) is the most frequently mutated gene in AML, mutated in approximately 30% of AML patients and over 50% of cytogenetically normal AML.[253,254] Under normal conditions NPM1 is a molecular chaperone, and has been implicated in facilitating centrosome duplication and ribosome biogenesis, as well as regulation of the ARF-p53 tumour suppressor pathway.[253] In order to complete these many roles NPM1 shuttles between the nucleus and cytoplasm but is predominantly found in the nucleus. When mutated in AML patients, the nuclear localisation signal is disrupted leading to an accumulation of NPM1 in the cytoplasm (NPM1c).[255] These mutations are always heterozygous, as a WT copy of NPM1 is required for cell survival.[256] The most common type of NPM1 mutation is type A, this mutation is a 4 nucleotide duplication in exon 12, causing a shift in the reading frame, replacing the last 7 amino acids of the C-terminus with 11 different amino acids, thereby altering the nuclear localisation signal into a nuclear export signal.[253]

A NPM1 mutant mouse has been created by Vassiliou et al. (2011).[240] This mouse is a conditional humanised Npm1c knock-in, with a type A NPM1 mutation induced by pIpC treatment (Mx1-Cre) (hereafter this mouse will be referred to as NPM1 mutant).

This NPM1 mutant mouse model has an increase in $Mac1^+$ and $Gr1^+$ cells, and reduction in B cells ($B220^+/CD19^+$), indicating a myeloid expansion/bias. NPM1 mutant HSC have increased serial re-plating capacity in methylcellulose, a surrogate for self-renewal potential. 30% of NPM1 mutant mice develop AML, leading to on average shorter life expectancy of NPM1 mutant vs WT animals. The mice that developed AML had enlarged spleens and livers, AMLs from these mice were transplantable. Additional mutations (induced by transposon mutagenesis using sleeping beauty transposon) accelerate AML development, with 80% animals developing AML within 1 year, indicating that additional mutations may be required to

combine with NPM1 (most common co-mutations in patients are FLT3-ITD and DNMT3A).

In vitro, HSCs from NPM1 mutant mice had slower entry into cell cycle, compared to WT HSCs (Figure 13 left, after 2 days in culture 87.03% of WT HSCs had divided vs 70.76% of NPM1 mutant HSCs, p=0.0002), and longer time to subsequent cell divisions (Figure 13 middle, after 3 days in culture 88.17% of WT HSCs had completed 2+ divisions, vs 62.64% of NPM1 HSCs, p=0.0001). This delay in cell division resulted in smaller average clone size in NPM1 mutant clones after 10 days in culture (Figure 13 right, VS, M, L p=0.0001, S p=0.021).

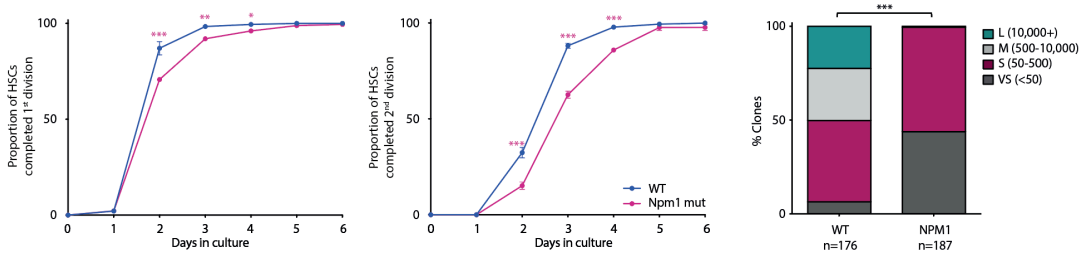


Figure 13 NPM1 HSCs have delayed cell division kinetics compared when with WT HSCs.
When compared with WT HSCs (blue line, N=2), a smaller proportion of NPM1 mutant HSCs (pink line, N=3) have entered into cell cycle after 2 in culture (p=0.0002) or completed 2+ cell divisions after 3 day in culture (p=0.0001). After 10 days in culture NPM1 HSCs had given rise to a larger proportion of VS clones than WT HSC (p=0.0001). Error bars are SEM.

3.2.4 P53 null HSCs have equivalent cell division kinetics to WT HSCs *in vitro*

P53 is a transcription factor which, when induced by stress, results in cell cycle arrest, apoptosis and senescence.[257] *TP53*, the gene that encodes p53 in humans, is the most frequently mutated gene in cancers, it is mutated or deleted in approximately 50% of all cancers including haematological malignancies.[258] 10% of primary AML and MDS patients have p53 mutations and 30% of patients who have MDS or AML secondary to radiation or chemotherapy have p53 mutations. P53 mutations are also commonly found in individuals with ARCH.[259]

A p53 knock-out mouse has been developed. This mouse has a 450 base pair deletion of the *Trp53* gene (mouse p53 gene), including the exon 5 splice acceptor site, this leads

to complete knock-out of p53 protein.[241] 74% of mice homozygous for this mutation developed tumours by 6 months of age, most commonly malignant lymphomas.

In HSCs, p53 has roles in maintaining quiescence and DNA damage response. P53 null mice have been reported to have an increased number of immunophenotypic HSCs (both LSKCD34$^-$ cells[260] and CD150$^+$Sca1$^+$c-Kit$^+$CD41$^-$CD48$^-$Lin$^-$ cells[261]), as p53 has roles in DNA damage response and apoptosis it has been suggested that this increase is due to decreased apoptosis from this population. However, Annexin-V staining of LSK fraction found no significant differences in levels of apoptosis compared with WT cells,[261] suggesting that an increase in self-renewal may be responsible for the increase in phenotypic HSCs. This is supported by serial competitive transplantation data that has shown p53 null HSCs have an enhanced self-renewal potential.[258]

The method by which loss of p53 achieves increased self-renewal has been hypothesised to be through its interaction with EZH2, the epigenetic regulator/polycomb protein commonly mutated in B-cell lymphomas, MDS, MPN, and T-cell ALL (see introduction section 1.10.2).[216] Overexpression of EZH2 in HSCs increases their self-renewal capacity. Mutant p53 acts to enhance EZH2s association with the chromatin, thus increasing the levels of H3K27me3 methylation and thereby influencing the expression of self-renewal and differentiation genes.[258]

When single p53 null HSCs were cultured *in vitro* they performed very similarly to WT HSCs in terms of entry into cell cycle (Figure 14 left, day 2 p=0.662), time to

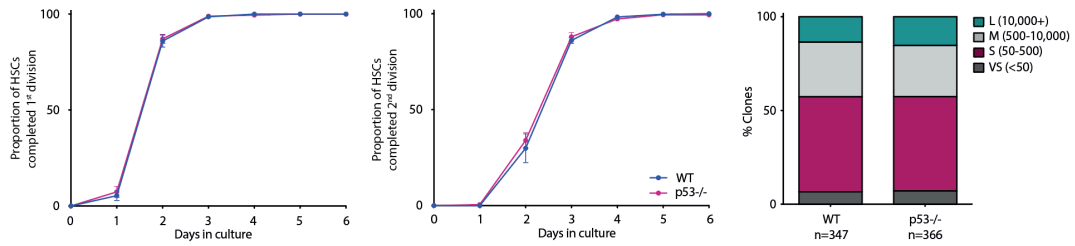


Figure 14 P53 HSCs behave similarly to WT HSCs in in vitro cultures
P53 HSC (pink line, N=4) have similar cell division kinetics to WT HSC (blue line, N=4) in terms of their time to first and second division (no statistical difference between proportion completed first division at day 2, p=0.662, or second division at day 3, p=0.505). P53 HSCs gave rise to a similar proportion of VS (p=0.883), S (p=0.940), M (p=0.618), and L (p=0.524) clones to WT HSCs after 10 days in culture. Error bars are SEM.

subsequent cell division (Figure 14 middle, day 3 p=0.505), and clone size after 10 days in culture (Figure 14 right, VS p=0.883, S p=0.940, M p=0.618, L p=0.524).

3.3 MPN patient JAK2 V617F and TET2 mutant HSPCs validate murine *in vitro* findings

To validate the findings from the JAK2 V617F and TET2 knock-out mouse models in patient samples, data was utilised from isolated human $Lin^{-}CD34^{+}CD38^{-}CD90^{+}CD45RA^{-}$ cells (hHSPCs) from MPN patients with JAK2 and TET2 mutations.[190] Single hHSPCs were cultured in stem cell maintenance conditions and divisional kinetics were tracked for 10 days. Individual clones were plated in a colony forming cell assay for a further 14 days before being harvested for genotyping (Figure 15), allowing retrospective assignment of mutational status to the initial sorted cells.

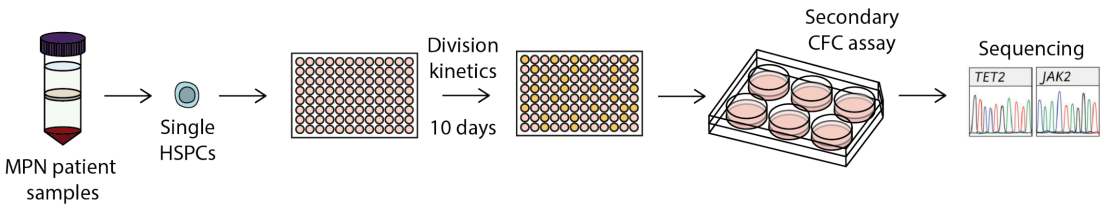


Figure 15 Schematic of in vitro cultures from patient samples
Single $Lin^{-}CD34^{+}CD38^{-}CD90^{+}CD45RA^{-}$ cells (HSPCs) were sorted from MPN patient peripheral blood samples (3 essential thrombocythemia, 5 polycythemia vera, 4 secondary myelofibrosis) into 96-well plates, and cell counts were performed daily for 10 days. On day 10 clones were harvested and plated in a colony-forming cell assay, and the colonies produced were harvested for genotyping.

Like in the mouse model, hHSPCs with a JAK2 mutation alone had a shorter time to entry into cell cycle (Figure 17 left, after 3 days in culture 74% of JAK2 mutant hHSPCs had divided vs 51.85% of non-mutant cells, p=0.015), and reduced time to subsequent cell division (Figure 17 middle, after 4 days 77.45% JAK2 mutant completion 2+ divisions vs 50.88% of non-mutant cells, p<0.0001). However, unlike in the mouse model, this did not lead to significantly larger clones after 10 days in culture (Figure 17 right, VS p=0.202, S p=0.786, M p=0.054, L p=0.674).

Like the TET2 mouse model, hHSPCs with a TET2 mutation alone did not have division kinetics significantly different to WT HSPCs (Figure 16 left and middle, 1st division at day 3 p=0.582, 2nd division at day 4 p=1). However, unlike the mouse model,

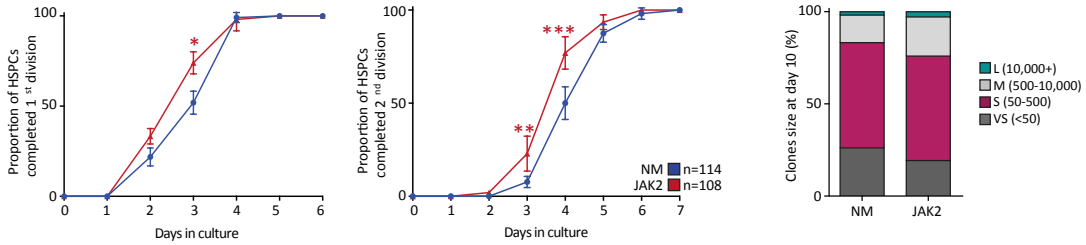


Figure 17 JAK2 V617F hHSPCs have accelerated cell division kinetics but do not make larger clones after 10 days in culture

After 3 days in culture the proportion of JAK2 V617F HSPCs (JAK2, red line, N=11) that have entered into cell cycle is higher than that in non-mutant HSPCs (NM, blue line, N=11) (p=0.015), and after 4 days in culture a larger proportion have completed 2+ cell divisions (p<0.0001). This did not lead to a larger proportion of large clones after 10 days in culture (p=0.674). Error bars are SEM.

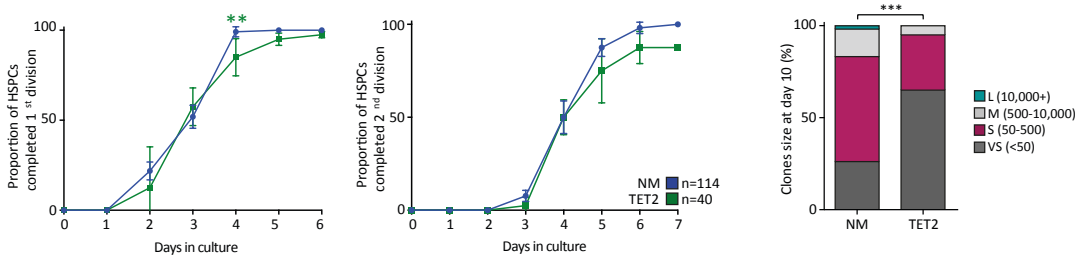


Figure 16 TET2 mutant hHSPCs have 'normal' cell division kinetics but make on average smaller clones after 10 days in culture

After 3 days in culture a similar proportion of TET2 mutant (TET2, green, N=2) and non-mutant hHSPCs (NM, blue, N=11) has entered into cell cycle (p=0.582), and after 4 days in culture a similar proportion had completed 2+ cell divisions (p=1). After 10 days in culture, compared with non-mutant clones, a larger proportion of TET2 mutant clones were very small (p<0.0001). Error bars are SEM.

this did result in on average smaller colonies after 10 days in culture (Figure 16 right, 65% of TET mutant colonies contain fewer than 50 cells, compared with 26.32% of non-mutant colonies, p<0.0001).

Like in the mouse models, JAK2 TET2 double mutant cells from patients closely resembled JAK2 V617F alone. With accelerated time to first division (Figure 18 left, 70.45% JAK2 TET2 double mutant cells divided by day 3 vs 51.85% of non-mutant cells, p=0.002) and second division (Figure 18 middle, 66.46% of JAK2 TET2 mutant cells completed 2+ divisions by day 4 vs 50.88% of non-mutant cells, p=0.013). Like with JAK2 mutation alone, no significant differences were observed in terms of day 10 colony size (Figure 18 right, VS p=0.120, S p=0.332, M p=0.588, L=1).

Despite differences in the day 10 colony size, the human data mirrors the division kinetics seen in the mouse models; JAK2 V617F increases proliferation of primary HSCs, in the presence or absence of a TET2 mutation, and that TET2 loss on its own does not confer a proliferation advantage on HSCs *in vitro*.

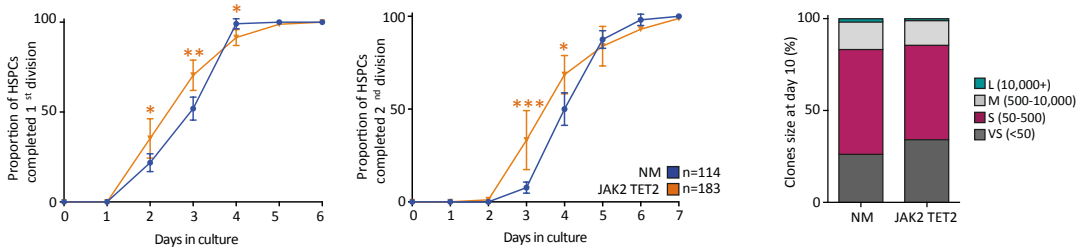


Figure 18 hHSPCs with JAK2 and TET2 mutations have accelerated cell division kinetics but do not make larger clones after 10 days in culture

After 3 days in culture the proportion of JAK2 TET2 double mutant HSPCs (JAK2 TET2, orange, N=8) that had entered into cell cycle was higher than that from non-mutant HSPCs (NM, blue, N=11)(p=0.002), and after 4 days in culture a larger proportion had completed 2+ cell divisions (p=0.013). This did not lead to a larger proportion of large clones after 10 days in culture (p=1). Error bars are SEM.

Results

4

MPN HSCs *in vivo*

In this chapter I explored the *in vivo* myeloproliferative and self-renewal capacity of JAK2 and TET2 mutant BM. Firstly I tested whether BM from JAK2 TET2 double mutant animals is able to impart an MPN-like phenotype in non-competitive serial transplantation recipients. Followed by establishing the minimum number of cells required to initiate this phenotype by limiting dilution. And finally, assessment of the self-renewal potential of JAK2 mutant, TET2 mutant, and double mutant BM by competitive serial transplantation. The additional transplantation data that was added for statistical comparison in section 4.2.2 was from an experiment performed by Juan Li.

4.1 Non-competitive transplantation

4.1.1 Primary recipients of JAK HOM TET HOM bone marrow display MPN-like phenotype in non-competitive transplantation setting

To assess whether JAK HOM TET HOM BM cells are capable of imparting any haematopoietic phenotype upon recipients, non-competitive transplantation was performed. 5x105 whole BM cells from a pool of 2 JAK HOM TET HOM (CD45.2) or 3 WT (CD45.2) mice were transplanted into sub-lethally irradiated W41/CD45.1 recipients. Mice were serially bled to assess chimerism and blood parameters (Figure 19).

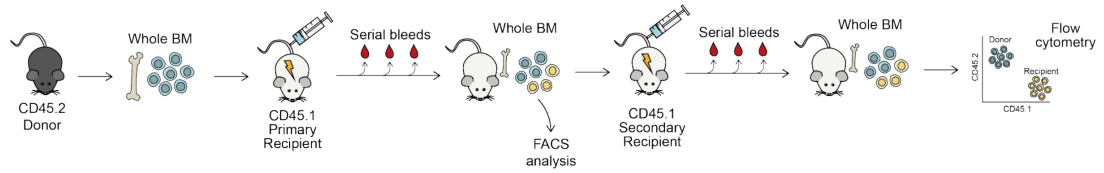


Figure 19 Schematic of non-competitive transplantation
Whole BM from JAK HOM TET HOM and WT donors (both CD45.2) was harvested and 5x10^5 cells transplanted into sub-lethally irradiated W41/CD45.1 recipients. Recipients were serially bled to assess for blood parameters and chimerism. 28 weeks after primary transplantation BM was harvest, a portion (1x10^7 cells) was secondarily transplanted into sub-lethally irradiated W41/CD45.1 recipients, while the remainder was analysed by flow cytometry for chimerism, mature lineage output, and HSC number. Secondary recipients were also serially bled and 18 weeks post-transplantation BM was harvested, and later analysed by flow cytometry. Donor chimerism was calculated as Donor/(Donor + Recipient).

Throughout the experimental time course, primary recipients of JAK HOM TET HOM mutant BM displayed similar levels of chimerism to WT BM recipients (Figure 20). In WT recipients chimerism grew from 65.32% at 8 weeks to 88.38% at 24 weeks post-transplantation, and in JAK HOM TET HOM recipients growing from 68.54% (p=0.257) at 8 weeks to 88.17% (p=0.949) at 24 weeks post-transplantation.

Mice expressing homozygous human JAK2 V617F have previously been shown to have a myeloproliferative phenotype resembling that seen in MPN PV patients[228] (PV

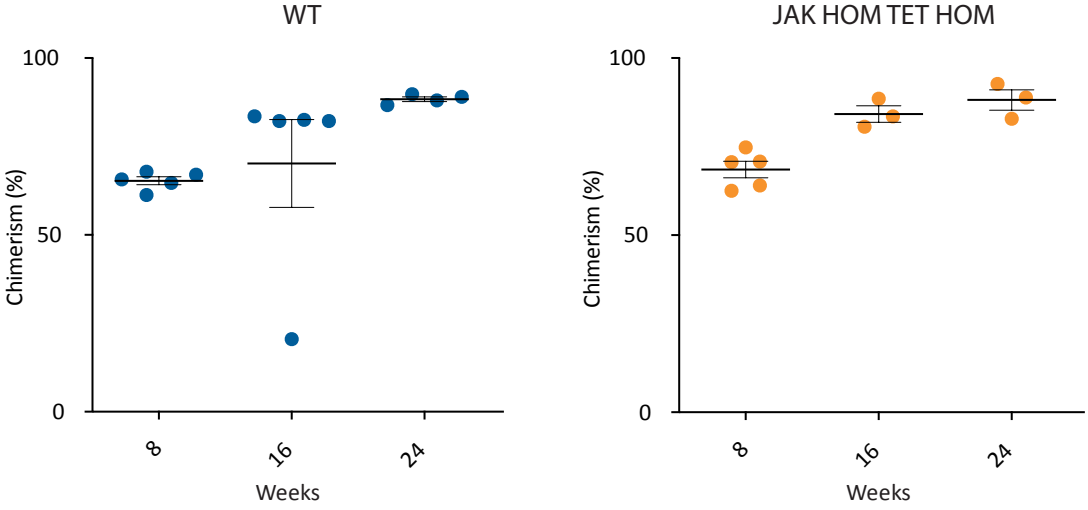


Figure 20 Donor chimerism in WT and JAK HOM TET HOM BM primary recipients in non-competitive transplantation
Chimerism was comparable between WT (left) and JAK HOM TET HOM (right) recipients at all time points (8 weeks p=0.257, 16 weeks p=0.326, 24 weeks p=0.949). Error bars are SEM.

covered in intro section 1.9.1); most simply defined by a strong erythrocytosis, measured by elevated haematocrit and haemoglobin. All blood parameters were measured in post-transplantation bleeds but only haematocrit and haemoglobin will be discussed in this thesis as these are the relevant parameters for establishing the PV-like phenotype.

Compared with recipients of WT BM, mice receiving JAK HOM TET HOM BM had elevated haematocrit and haemoglobin throughout the experiment (Figure 21). WT haematocrit remained relatively stable (at 16 weeks average HCT 48.16%) whereas JAK HOM TET HOM recipient haematocrit was significantly higher (at 16 weeks average HCT of 92.62%, p=0.019). Similarly, the haemoglobin levels of WT recipients were largely stable (at 16 weeks, average HGB of 15.32g/dL), while the haemoglobin of JAK HOM TET HOM recipients was significantly elevated at all time points (at 16 weeks average HGB of 26.60g/dL, p=0.026).

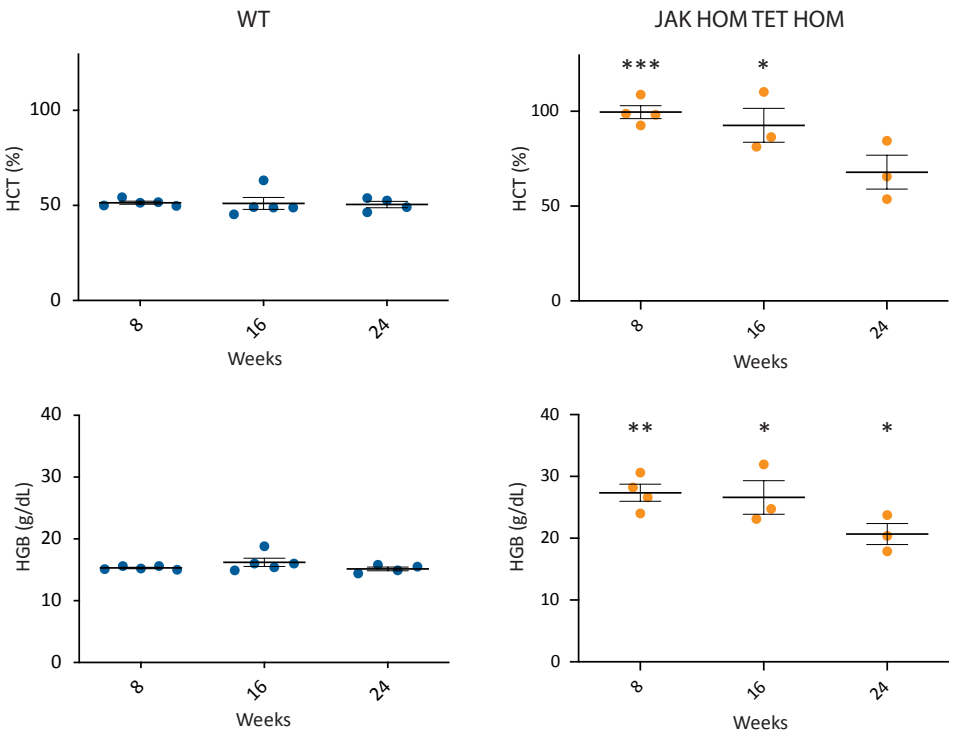


Figure 21 Haematocrit and haemoglobin in primary non-competitive transplantation recipients of JAK HOM TET HOM bone marrow
Recipients of WT BM in blue (left) and recipients of JAK HOM TET HOM in orange (right). JAK HOM TET HOM BM recipients display elevated HCT (p=0.019) and HGB (p=0.026, both at 16 weeks post-transplantation). Error bars are SEM.

From the primary transplantation data, it can be concluded that JAK HOM TET HOM repopulation capacity is equivalent to that of WT animals, and JAK HOM TET HOM cells are able to impart a myeloproliferative phenotype, as evidenced by elevated haemoglobin, for at least 24 weeks.

28 weeks after primary transplantation, recipient mice were sacrificed and $1x10^7$ whole BM cells (containing recipient and donor-derived cells) were secondarily transplanted into sub-lethally irradiated W41/CD45.1 recipients. The remaining BM was analysed by flow cytometry. Fully differentiated erythrocytes do not read out in the chimerism measurement, as they lack the CD45 antigen, by also considering the chimerism within the BM this lineage is able to contribute to chimerism as erythroid progenitors still possess CD45. Although this does report the full extent of erythroid chimerism it allows the erythroid lineage to influence the chimerism to some extent. Like in the peripheral blood, WT and JAK HOM TET HOM BM recipients had comparatively high chimerism in the BM (WT 94.70%, JAK HOM TET HOM 97.30%, p=0.359) (Figure 22 left).

The lineage composition of the donor populations were assessed by flow cytometry, no significant differences were seen in the proportion of the graft made up of granulocytes and macrophages, B-cells, T-cells, or other cell types (Figure 22 middle, GM p=0.459, B-cells p=0.407, T-cells p=0.704, other p=0.678). The proportion of the donor population which are immunophenotypically defined as HSCs (ESLAM) was not significantly different between WT and JAK HOM TET HOM recipients (Figure 22

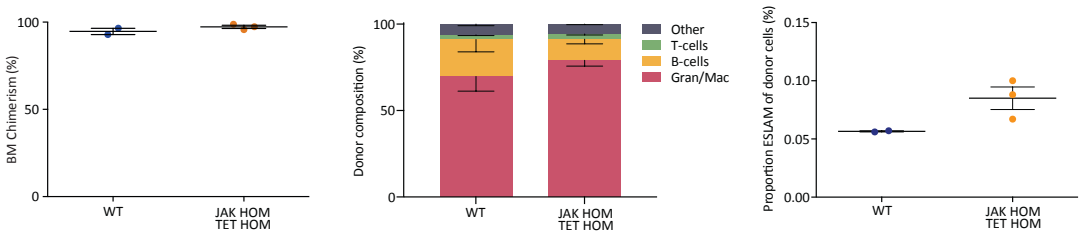


Figure 22 Chimerism and bone marrow composition in bone marrow of non-competitive primary transplantation recipients.
JAK HOM TET HOM and WT BM recipients have eqivalent chimerism in primary BM (left), p=0.359, and no significant differences in donor graft lineage composition (centre) (GM (Mac1$^+$, Gr1$^{+/-}$) p=0.459, B-cells (B220$^+$) p=0.407, T-cells (CD3$^+$) p=0.704, other p=0.678 (Mac1$^-$, Gr1$^-$, CD3$^-$, B220$^-$)) or proportion of ESLAMs (right), p=0.098. Error bars are SEM.

right, WT average 0.057, JAK HOM TET HOM average 0.085, p=0.098). While the assessment of ESLAM cells post-transplantation gives a representation of HSCs numbers, it should be noted that the ESLAM sorting strategy has not been fully validated in the post-transplantation setting, indeed single cell secondary transplantation of ESLAM cells has shown a HSC purity of ~30% in this population.[262]

4.1.2 JAK HOM TET HOM BM recipients continue to display MPN-like phenotype in secondary non-competitive transplantation

In order to determine the longevity of the phenotype seen in primary transplantation, secondary transplantation was performed. As in the primary transplantation, the secondarily transplanted mice were serially bled to assess for chimerism and blood parameters.

8 weeks post-secondary transplantation, chimerism remained comparable between WT and JAK HOM TET HOM recipients (Figure 23 top, 8 weeks WT 46.05%, JAK HOM TET HOM 41.73%, p=0.834) , and JAK HOM TET HOM recipients continuing to display elevated haematocrit (Figure 23 middle, at 8 weeks WT 57.30%, JAK HOM TET HOM 94.45%, p=0.053) and haemoglobin (Figure 23 bottom, 8 weeks WT 14.85g/dL, JAK HOM TET HOM 23.99g/dL, p=0.048). Two recipients of JAK HOM TET HOM BM died of unrelated causes between the 8- and 16-weeks post-transplantation, removing the power to draw statistically significant findings from this final timepoint, however the one remaining JAK HOM TET HOM recipient continued to display strong chimerism, which was accompanied by elevated haematocrit and haemoglobin (Figure 23).

18 weeks after secondary transplantation mice were sacrificed, BM was frozen and later analysed by flow cytometry. As in the blood, chimerism levels were comparable between JAK HOM TET HOM and WT recipients (WT1 91.4%, WT2 94.1%, JAK HOM TET HOM 93.6%) (Figure 24 left). The proportion of the graft which is comprised of immunophenotypic HSCs (ESLAM) for both genotypes seems to be lower than after the primary transplantation (Figure 24 right), however it is difficult to conclusively state weather both genotypes have comparable proportions due to the unrelated animal deaths causing low replicate numbers for the JAK HOM TET HOM

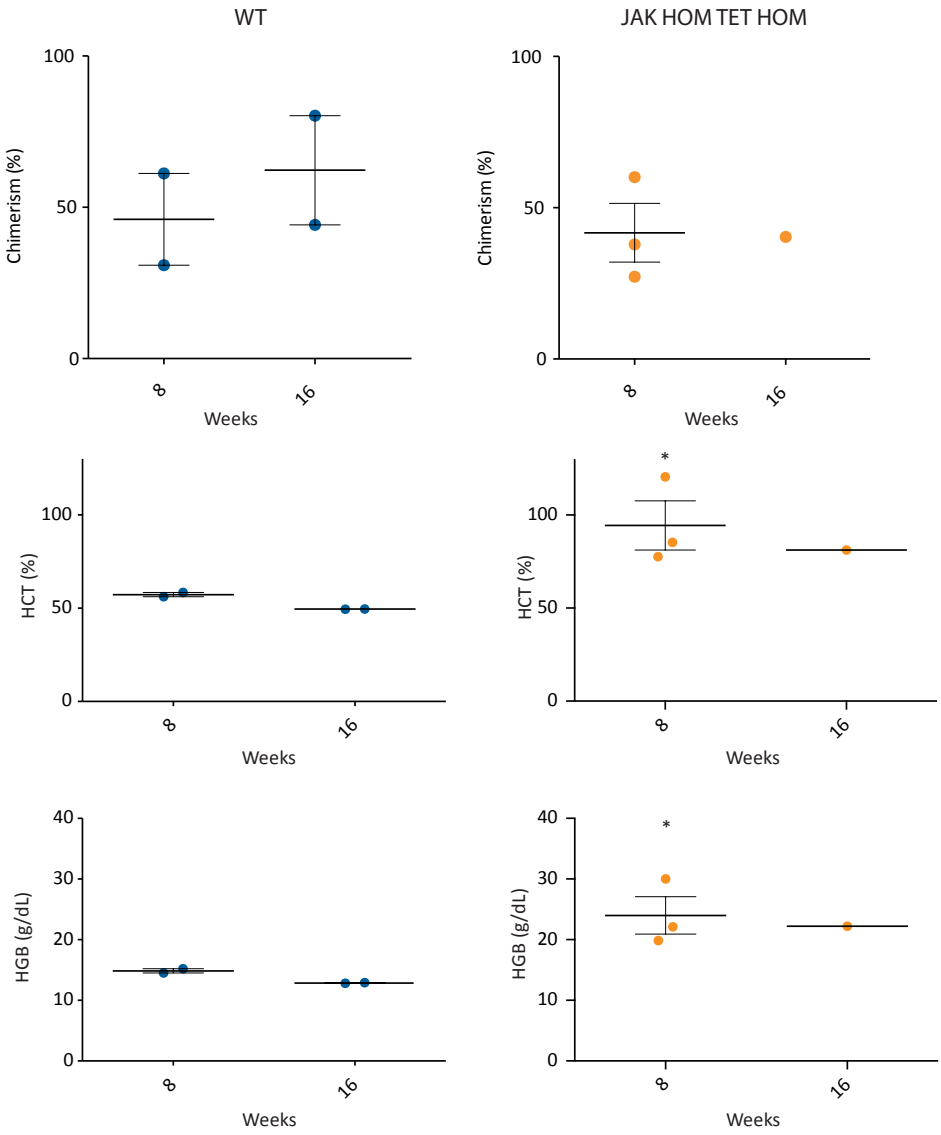


Figure 23 Chimerism and blood parameters of secondary recipients of non-competitively transplanted WT and JAK HOM TET HOM bone marrow
At 8 weeks post-transplantation no difference in chimerism between WT and JAK HOM TET HOM recipients (p=0.834), and a strong erythrocytic phenotype in JAK HOM TET HOM recipients (HCT p=0.053, HGB p=0.048).

arm. Likewise, no conclusive statements can be drawn about the balance of linages present in the BM (Figure 24 middle). The freezing of the BM before analysis may have also influenced the balance of cell types, as some types may be more sensitive to freezing and thawing than others.

From this set of transplantations, it can be concluded that in the non-competitive setting JAK HOM TET HOM BM is able to engraft in recipient animals and impart a myeloproliferative phenotype in both primary and secondary transplantation recipients,

as evidenced by elevated haematocrit and haemoglobin, indicating a significant expansion in the erythroid population.

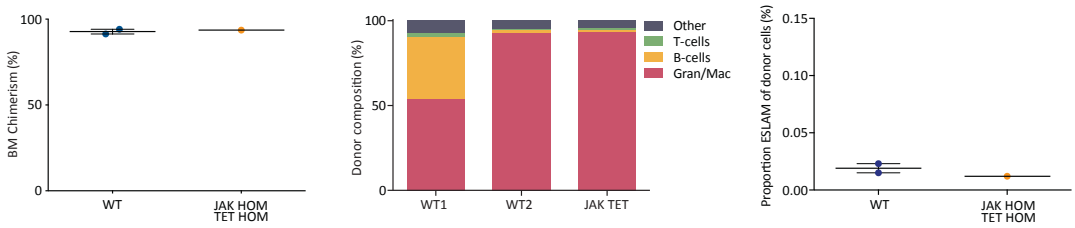


Figure 24 Chimerism and bone marrow composition in non-competitive secondary transplantation bone marrow
Due to mouse numbers statistical test cannot be performed, however it seems that JAK HOM TET HOM and WT BM recipients have eqivalent chimerism in secondary BM (left), similar donor graft lineage composition (centre), and proportion of ESLAMs (right). For WT error bars are SEM.

4.2 50,000 JAK HOM TET HOM cells or more are able to engraft and generate MPN-like phenotype upon transplantation

In order to determine the minimum number of JAK HOM TET HOM cells required to initiate a myeloproliferative phenotype, a limiting dilution transplantation assay was performed by competing 100,000, 50,000 or 10,000 WT or JAK HOM TET HOM BM cells against 100,000 WT competitor cell. Recipients were serially bled to assess for an MPN-like phenotype (Figure 25).

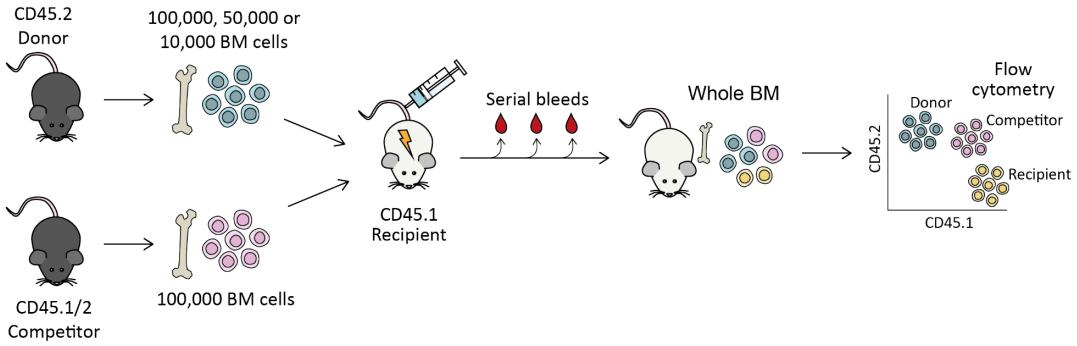


Figure 25 Schematic of limiting dilution competitive transplantation
100,000, 50,000 or 10,000 WT or JAK HOM TET HOM BM cells from CD45.2 donors were transplanted, alongside 100,000 WT competitor cell (CD45.1/.2), into W41/CD45.1 sub-lethally irradiated recipients. Recipients were serially bled to assess for blood parameters and chimerism. Donor chimerism was calculated as a proportion of transplanted cells (i.e. Donor/(Donor + Competitor)).

Peripheral blood (PB) chimerism levels were comparable between WT and JAK HOM TET HOM recipient, with both showing a cell dose dependent chimerism. 10,000 BM cells from either WT or JAK HOM TET HOM donors was insufficient to give sustained chimerism out to the final timepoint, with both falling to <1% by 24 weeks post-transplantation (Figure 26 top, WT 0.94%, JAK HOM TET HOM 0.32%, no p as only 1 animal per group). In the 50,000 cells group, both WT and JAK HOM TET HOM started with similar levels of chimerism, at 4 weeks WT average chimerism was 23.1% and JAK HOM TET HOM was 26.4%. However, WT chimerism expanded more through the subsequent time points with final average chimerism at 24 weeks for WT recipients was 59.5%, while JAK HOM TET HOM average chimerism was 35.8%. As there is only 1 mouse in each arm it is not possible to conclude if this is a consistent difference; however, it alludes to the possibility of a slight self-renewal disadvantage in JAK HOM TET HOM HSCs. For the highest dose, 100,000 cells, WT and JAK HOM TET HOM were once again similar in their levels of chimerism (starting at 41.6% for WT and 50.4% for JAK HOM TET HOM and rising to 71.3% for WT and 84.6% for JAK HOM TET HOM).

Due to the low chimerism, 10,000 JAK HOM TET HOM cells did not cause a strong MPN-like phenotype in the recipient (Figure 26 middle and bottom, at 16 weeks post-transplantation HCT 69.7%, HGB 16.2g/dL, average WT HCT 60.74%, HGB 14.7g/dL). Higher doses of JAK HOM TET HOM BM (50,000 and 100,000) were able to impart a robust myeloproliferative phenotype upon recipients (at 16 weeks 50,000 JAK HOM TET HOM HCT 80%, HGB 19.2g/dL, 100,000 JAK HOM TET HOM HCT 81.92% HGB 19.8g/dL). Given that in JAK HOM TET HOM recipients chimerism between ~30 and 85% is accompanied with the same extent of HCT and HGB elevation, this could indicate that there is a maximum HCT and HGB able to develop in the recipient mice.

In non-transplanted JAK HOM mice average haematocrit is 80%-90%, and haemoglobin is 22-25g/dL.[228] The haematocrit from the limiting dilution transplantation is within the non-transplanted range (over 80%), however the HGB is lower (over 19g/dL). This difference in the maximum extent to which haemoglobin is elevated could be due to a number of reasons. Firstly, the non-transplanted mice JAK2

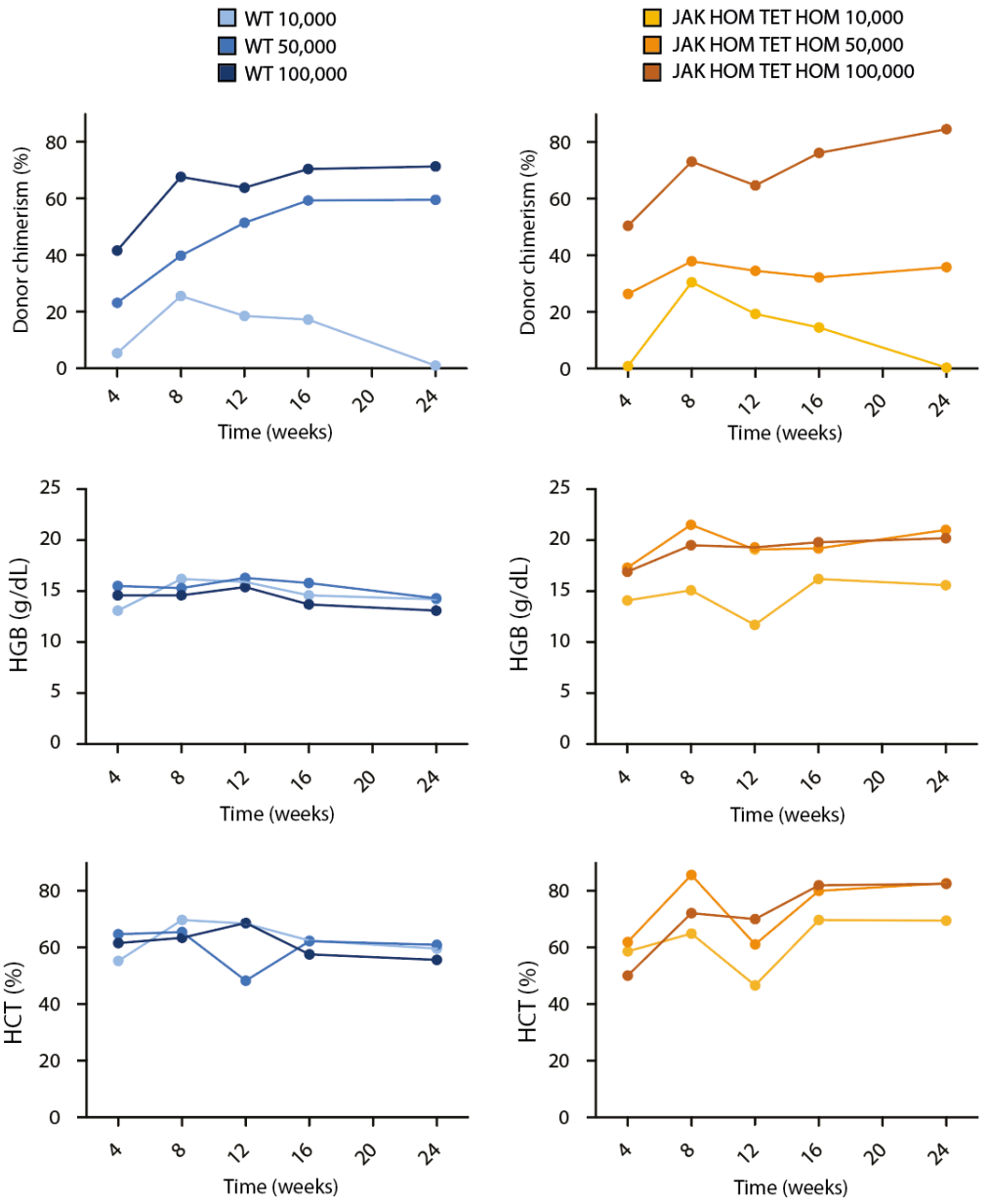


Figure 26 Chimerism and blood parameters of WT and JAK HOM TET HOM limiting dilution transplantation
Chimerism was not different between recipients of WT (left) and JAK HOM TET HOM BM (right). Transplantation of 50,000 or 100,000 JAK HOM TET HOM cells resulted in HCT and HGB above normal levels.

V617F is expressed from the foetal stage, whereas in the transplantation the mutation is introduced at time of transplantation, the earlier introduction might allow higher haemoglobin to develop over time. Secondly, in the steady state mouse the recombined allele is present in all cell types, not just haematopoietic cells, expression of JAK2 V617F in other cells within the BM microenvironment might play a role in this increase

in HGB. The BM microenvironment also could be affected by a number of other factors, including irradiation and room cleanliness, which could play a role in feedback and consequently haemoglobin. Thirdly, loss of TET2 could be reducing the extent to which haemoglobin can be elevated (however this seems unlikely as HCT is elevated to same extent).

From these findings it can be concluded at cell doses of 50,000 JAK HOM TET HOM cells or more able to engraft and generate MPN-like phenotype. Using limiting dilution analysis, an estimate of the disease-initiating cell frequency can be calculated for double-mutant whole BM (using WEHI eLDA); ~1 in 33 000 cells in the double mutant BM is a disease initiating cell, a number similar to the expected long-term HSC frequency in normal BM (1 in 20 000).

4.3 Competitive whole bone marrow transplantation from JAK2 and TET2 mutant mice

As mentioned in the introduction, serial competitive transplantation is the gold standard assay for determining the self-renewal capacity of HSCs (section 1.6.2). Competitive transplantation also represents a more disease-like state, where the leukaemic clone must compete against non-mutant cells to drive disease. To determine the relative ability of BM cells from TET2 mutant, JAK2 mutant, and double mutant mice to compete with non-mutant cells and initiate an MPN phenotype in recipients, and the

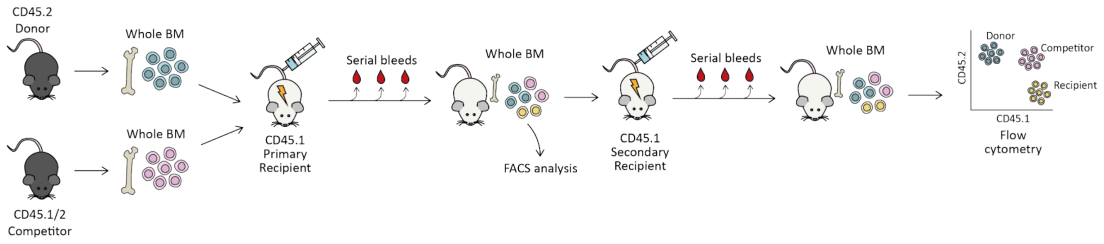


Figure 27 Schematic of serial competitive transplantation
$5x10^5$ WT or mutant BM cells from CD45.2 donors were transplanted, alongside $5x10^5$ WT competitor cell (CD45.1/.2), into W41/CD45.1 sub-lethally irradiated recipients. Recipients were serially bled to assess for blood parameters and chimerism. 28 weeks post-transplantation mice were sacrificed and BM harvested, $1x10^7$ cells (containing donor, competitor and recipient cells) were secondarily transplanted and the remainder analysed by flow cytometry. Secondary transplantation recipients were serially bled to assess for MPN-like phenotype. Donor chimerism was calculated as a proportion of transplanted cells (i.e. donor/(donor + competitor)).

effect of each mutation on HSC self-renewal, serial competitive transplantation was undertaken. For this $5x10^5$ whole BM cells from pools of mutant or WT mice (WT 3 mice, TET HOM 3 mice, JAK HOM 3 mice, JAK HOM TET HET 2 mice, JAK HOM TET HOM 2 mice, all CD45.2) were transplanted, alongside $5x10^5$ whole BM cells from CD45.1/2 WT mice, into sub-lethally irradiated W41/CD45.1 recipients (Figure 27). Mice were serially bled to assess for MPN-like phenotype and chimerism. 28 weeks after primary transplantation mice were sacrificed and BM secondarily transplanted. The results of primary transplantation for each genotype will be reported first, followed by the results of secondary transplantation.

4.3.1 Loss of TET rescues JAK2 V617F self-renewal defect in primary transplantation

Transplantation of WT BM acted as a control and was used for statistical comparison with the mutant models. The donor chimerism of WT BM recipients grew throughout the experimental time course (Figure 28 left) from an average of 52.7% at 4 weeks to 73.5% at 24 weeks. These mice had normal blood parameters; at 16 weeks post transplantation haemoglobin of 14.7g/dL and haematocrit of 45.17% (Figure 28 right and middle). The WT recipients establish a normal range of haematocrit and haemoglobin for this experiment; HCT below 70, HGB below 18. Between the 16th and 24th week post-transplantation 2 of 3 mice were culled owing to reasons unrelated to the transplantation, consequently statistical tests at the 24-week time point were not possible.

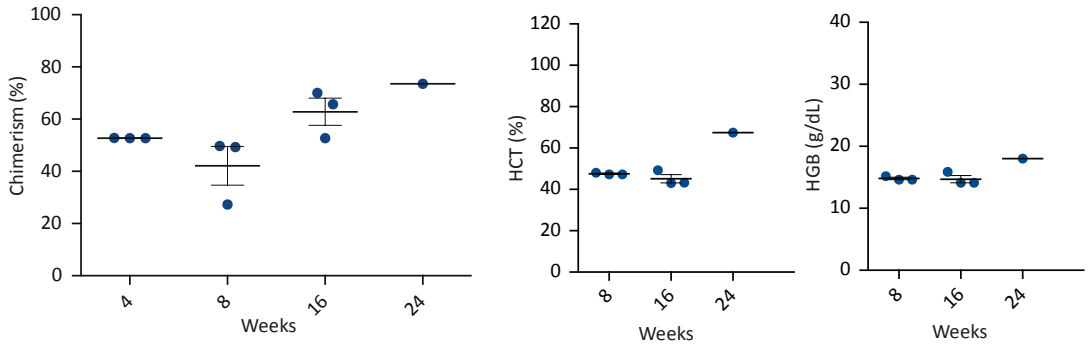


Figure 28 Chimerism and blood parameters from primary recipients of competitively transplantation of WT BM

Average chimerism (left), HCT (centre) and HGB (right) from recipient of WT BM. From this data the normal range of haematocrit and haemoglobin for transplantation recipients is established as HCT below 70% and HGB below 18g/dL.

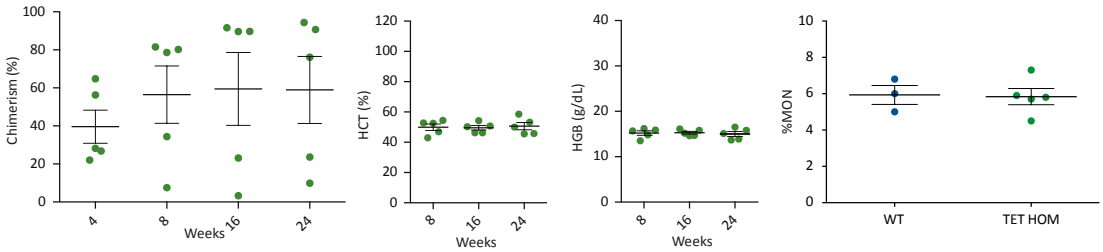


Figure 29 Chimerism and blood parameters from recipients of competitively transplanted TET HOM BM

Average chimerism of TET HOM recipients was similar to that of WT recipients (p=0.874 at 16 weeks post-transplantation). No changes in blood parameters were observed in TET HOM recipients; at 16 weeks HCT (p=0.077), HGB (p=0.225), and percentage monocytes (p=0.449) were all comparable to that of WT recipients. Error bars are SEM.

Primary recipients of TET HOM BM had an average donor chimerism similar to that of WT recipients (54.04% at 4 weeks to 58.94% at 24 weeks post-transplantation). Although the variability in TET HOM chimerism was much higher than in WT recipients, with 3 mice with rising chimerism throughout the time course (by 24 weeks over 75%), and the other 2 mice chimerism levels falling, (by 24 weeks below 25%) (Figure 29 far left).

Recipients of JAK HOM BM had an initially high chimerism at (Figure 30 left, 4 weeks post transplantation average chimerism 84.72%, p=0.0006) however this dropped off steeply and remained significantly low for the remaining time points (8 weeks 8.168%

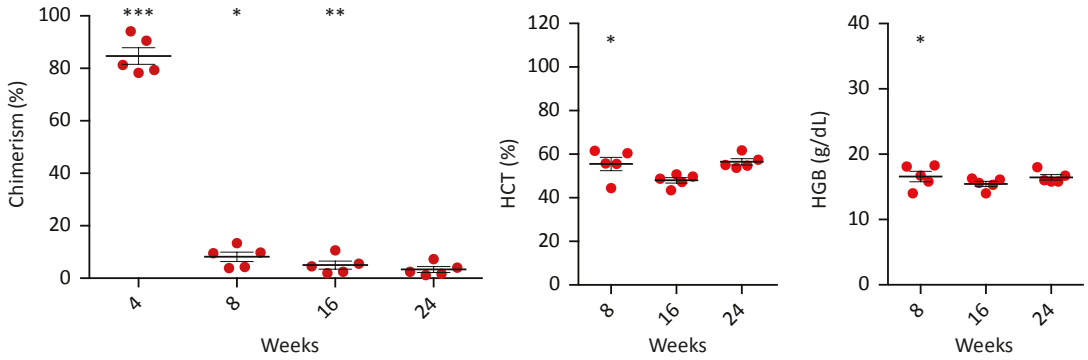


Figure 30 Chimerism and blood parameters of primary recipients of competitively transplanted JAK HOM bone marrow

Chimerism was initially high at 4 weeks post-transplantation (p=0.0006), and at 8 weeks was sufficient to produce elevated HCT (p=0.029) and HGB (p=0.044), however by 16 weeks HCT and HGB were not significantly different to normal (HGB p=0.178, HCT p=0.154). Error bars are SEM.

p=0.001, 16 weeks 5.02% p=1.08x10^{-5}, 24 weeks 3.34%). Due to these low levels of donor chimerism, an MPN-like phenotype was not observed in the JAK HOM recipients; haematocrit and haemoglobin were within the normal range (Figure 30 middle and right, at 16 weeks average HCT 48%, p=0.154, and HGB 15.46g/dL, p=0.178).

JAK HOM HSCs have been shown to be hyperproliferative *in vitro* (Chapter 3.1), this hyperproliferation is thought to lead to the initial burst in donor population size, but this hyperproliferation appears to come at the expense of HSC self-renewal, and the HSCs quickly exhaust, and donor population disappears as the cells differentiate and die.

Recipients of JAK HOM TET HET BM had an initially strong chimerism (Figure 31 left, 4 weeks average 58.42%) but this fell during the course of the primary transplantation to 20.72% at 24 weeks. Haematocrit and haemoglobin mirrored the level of chimerism; when chimerism was high in early time points HCT and HGB were elevated (Figure 31 middle and right, at 8 weeks when recipients had an average donor percentage of 33.54%, HGB 20.56g/dL p=0.002, HCT 64.56% p=0.011). As chimerism fell so did HCT and HGB; by 24 weeks, when chimerism was 20.716%, haemoglobin was 15.78g/dL, and haematocrit 57.80%. This is the upper cusp of the normal threshold. These results show that, while homozygous JAK2 V617F alone exhausts very quickly, combining a JAK2 mutation with heterozygous loss of TET2, in primary

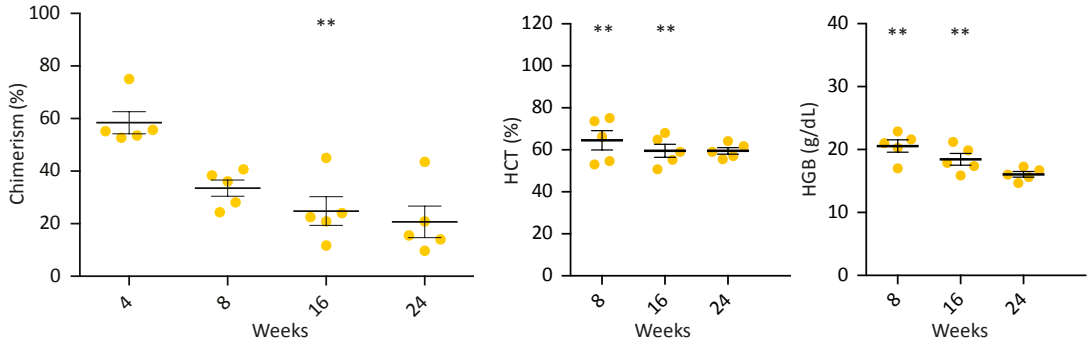


Figure 31 Chimerism and blood parameters of primary recipients of competitively transplanted JAK HOM TET HET bone marrow

JAK HOM TET HET donor chimerism fell throughout the time course; by 16 weeks post transplantation chimerism was significantly low (p=0.003), however HCT and HGB remained significantly high (HCT p=0.004, HGB p=0.008), showing the strong erythrocytic phenotype invoked by these cells.

transplantation, slows exhaustion leading to higher chimerism and, when chimerism is robust, an increase in HCT and HGB.

Unlike recipients of JAK HOM and JAK HOM TET HET BM, the chimerism of recipients of JAK HOM TET HOM BM did not fall but remained relatively stable throughout the primary transplantation (Figure 32 left, from 47.22% at 4 weeks, to 28.9% at 24 weeks). As a consequence of this sustained higher chimerism, haematocrit and haemoglobin were elevated throughout the primary transplantation (Figure 32 middle and right); at 16 weeks the average haemoglobin for JAK HOM TET HOM recipients was 25.52g/dL (p=0.003), and average haematocrit 84.26% (p=0.002).

When the models with a JAK2 mutation are compared, it can be concluded that JAK2 V617F alone has a strong self-renewal defect (leading to graft exhaustion), and while JAK HOM TET HET seems to slow the exhaustion of HSCs, JAK HOM TET HOM appears to, in primary recipients, rescue this exhaustion, with very little decrease in chimerism seen across the experimental time course. As a consequence of this robust chimerism HCT and HGB levels are very high across all time points. These data indicated a "dose" dependent response to loss of TET2, with loss of a single copy (JAK HOM TET HET) acting as a half-way response between no loss (JAK HOM alone) and complete loss of TET2 (JAK HOM TET HOM).

At 28 weeks post-transplantation, recipients were sacrificed and secondary transplantation performed. At this time chimerism and cellular composition of the BM

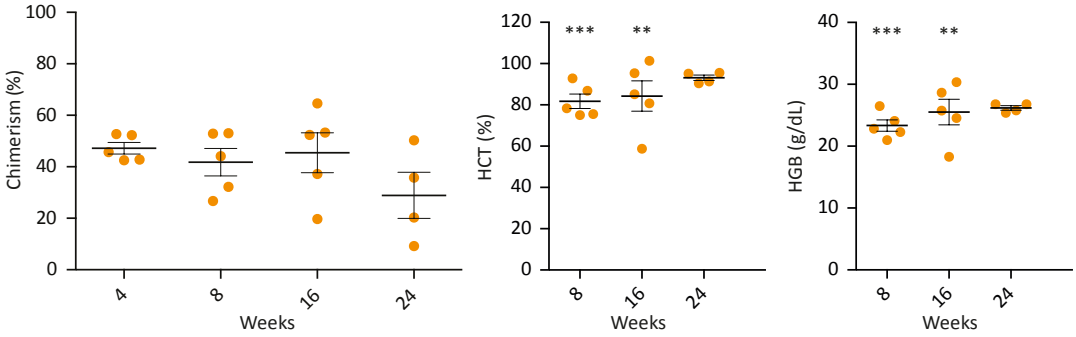


Figure 32 Chimerism and blood parameters of primary recipients of competitively transplanted JAK HOM TET HOM bone marrow

Chimerism remained relatively stable in primary transplantation recipients, this was accompanied by significantly elevated HCT (p=0.002, at 16 weeks) and HGB (p=0.003, at 16 weeks) compared with WT recipients. Error bars are SEM.

was assessed by flow cytometry. In the BM the broad trends in chimerism are similar to the PB (Figure 33 left); JAK HOM chimerism is virtually absent, TET HOM has robust chimerism with an average similar to WT. JAK HOM TET HET chimerism is low (below 25%) but unlike JAK2 alone all mice have chimerism above 1% (the cut off for positive repopulation). And JAK HOM TET HOM is positively repopulated for all with a much higher average chimerism than JAK HOM alone.

JAK HOM chimerism is very low in the BM of primary recipients (average 1.83%) only just over the cut off for positively repopulated, these grafts are comprised of a higher proportion of T- and B-cells than in WT/TET/JAK TET recipients (B cells comprise an average of 54.63% of the JAK HOM grafts vs 6.76% of JAK HOM TET HOM grafts, p=0.05) (Figure 33 middle). This is not representative of a "lymphoid bias" in JAK HOM HSCs, but is a result of the fact that these cell types have a longer half-life than myeloid cells and are therefore able to contribute to chimerism for longer, reading out even when the mutant stem and progenitor cells have stopped producing new cells.

The small graft size means conclusions about the graft composition have to be interpreted carefully, as the small denominator can make population proportions less reliable. For example, the proportion of ESLAM cells within the JAK HOM donor

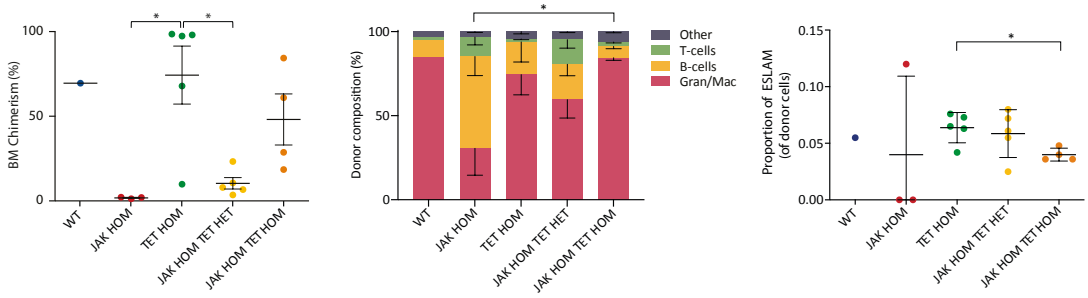


Figure 33 Chimerism and composition of the bone marrow of primary competitive transplantation recipients

Donor chimerism was significantly lower in JAK HOM (p=0.013) and JAK HOM TET HET (p=0.019) recipients relative to TET HOM recipients, JAK HOM TET HOM however was not significantly reduced (p=0.289). Comparison of the composition of donor grafts shows that JAK HOM has a significantly higher proportion on B-cells than JAK HOM TET HOM donor populations (p=0.05). The proportion of ESLAM HSCs within donor grafts was lower in JAK HOM TET HOM recipients than TET HOM recipients (p=0.013). Error bars are SEM.

population (Figure 33 right) for 2/3 mice is 0, but for the third mouse the proportion of phenotypic HSCs within the graft appears normal, the actual number of ESLAM cells within this graft is small, but as the total number of donor cells is also small a few more or less cells in the gate makes a huge difference to the proportion read out.

TET HOM recipients had higher chimerism in the BM than in the PB, this could be due to the BM timepoint being 4 weeks after the previous bleed, which could account for some of the difference, but as the chimerism had been fairly stable from 8 to 24 weeks this increase is more likely due to the difference in cellular composition of the BM. In agreement with literature, for all genotypes, the proportion of GM cells of donor origin was higher in the BM than the PB (Figure 34).[263] TET2 knock-out animals have been reported to have a GM bias (Ko et al. 2011)[210], this GM bias may account for the increased chimerism in the BM; if TET HOM HSCs produce proportionally more GM cells, are there are more GM cells resident in the BM than PB, it is reasonable that TET HOM would have a higher chimerism in the BM than the PB.

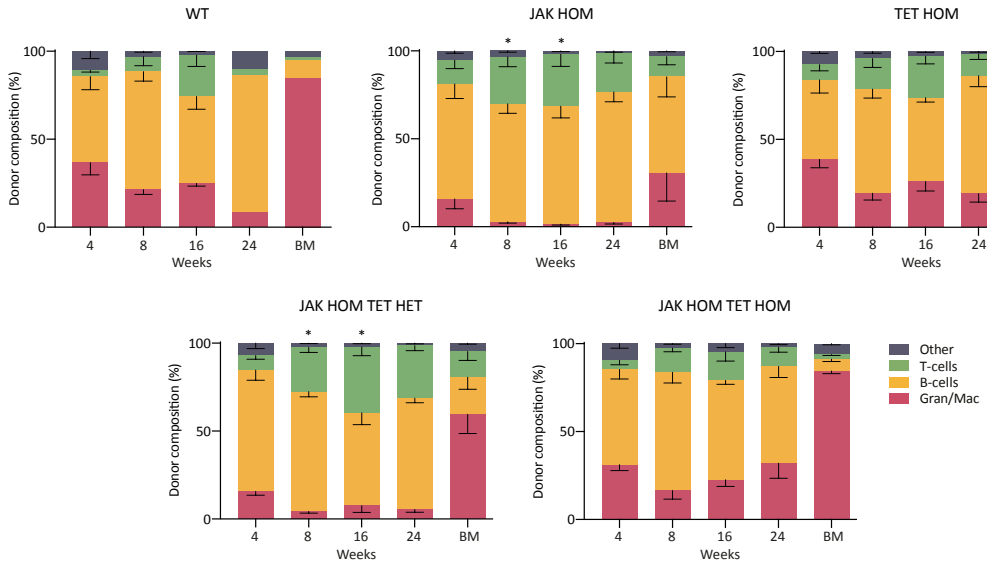


Figure 34 Donor graft composition in primary recipient peripheral blood and bone marrow for each genotype

In BM of all genotypes, granulocytes (Mac1$^+$, Gr1$^+$) and macrophages (Mac1$^+$, Gr1$^-$) made up a larger proportion of the donor population than in PB. In JAK HOM and JAK HOM TET HET recipients at 16 weeks post-transplantation the grafts had a reduced proportion of GM cells (JAK HOM p=0.0001, JAK HOM TET HET p=0.002), and increased proportion of lymphocytes (JAK HOM B-cells p=0.0146, JAK HOM TET HET T-cells p=0.044).

TET HOM has the highest average proportion of donor cells defined as ESLAM (Figure 33 significantly higher than JAK HOM TET HOM recipients, p=0.013), this high number of phenotypically defined HSCs is in keeping with literature reports of expansion of the stem cell pool and enhanced self-renewal of HSCs from TET knock-out mice.

BM analysis of JAK HOM TET HET recipients showed they had low average chimerism (10.42%, Figure 33), however all recipients were still positively repopulated (above 1%). A lower proportion of JAK HOM TET HET grafts were comprised of GM cells (not significant), this may hint towards exhaustion of the graft as GM cells have a shorter life. Of all the models with a JAK2 V617F mutation, JAK HOM TET HOM has the highest chimerism, and higher proportion of GM cells contributing to chimerism than JAK HOM TET HET, this implies that grafts are still actively contributing to production of shorter lived cell types. Expansion in 'other' cell types group, combined with low numbers of phenotypic HSCs, suggests a progenitor expansion in recipients of JAK HOM TET HOM BM. This progenitor expansion, in combination with mature cell expansion is thought to be the cause of the reduction in the proportion of phenotypic HSCs, not a reduction in the number of HSCs in these animals.

4.3.2 Secondary transplantation reveals self-renewal properties of HSCs with JAK2 and TET2 mutations

As the self-renewal defect in JAK2 HSCs is so strong, from the primary transplantation data it is clear loss of TET2 is able to at least partially rescue this defect, however self-renewal potential can only be definitively tested in secondary transplantation. To investigate the self-renewal potential of each genotype BM from primary transplanted recipients was harvested and $1x10^7$ cells (containing a mix of donor, competitor and recipient cells) was transplanted into sub-lethally irradiated W41/CD45.1 recipients. In the secondary transplantation many of the same effects were seen that were established in the primary transplantation recipients continue.

By the end of the primary transplantation there was only one remaining WT recipient from this experiment (Figure 36), due to 2/3 mice having been culled due to non-transplantation related health issues. In secondary transplantation this recipient continued to have HCT and HGB within the normal range.

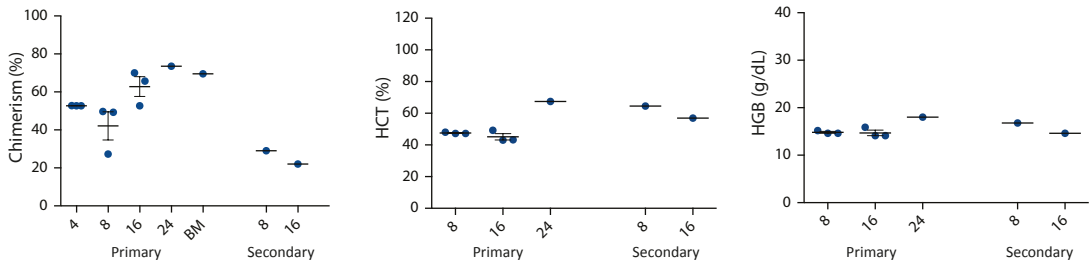


Figure 36 Chimerism and blood parameters of secondary recipient of competitively transplanted WT bone marrow
Secondary transplantation of the one remaining WT BM recipient continued to have HCT and HGB within normal range.

In order to be more confident in the conclusions drawn from this experiment, and allow statistical comparison of mutant genotypes with WT, WT recipients from another set of transplants were combined with this data set. The additional data was from a transplantation that was set up in the same way as this experiment ($5x10^5$ donor whole BM cells alongside $5x10^5$ competitor whole BM cells). The only common timepoint between these two experiments was the 16 week timepoint, this timepoint is therefore used for chimerism comparison.

Comparison of the primary 16 week post-transplantation chimerism data showed that the two WT data sets were not different (p=0.141) (Figure 35), and addition of this data did not alter the significance of JAK HOM when compared with WT (without extra data p=0.005, with extra data p=0.006). However, addition did alter the degree of

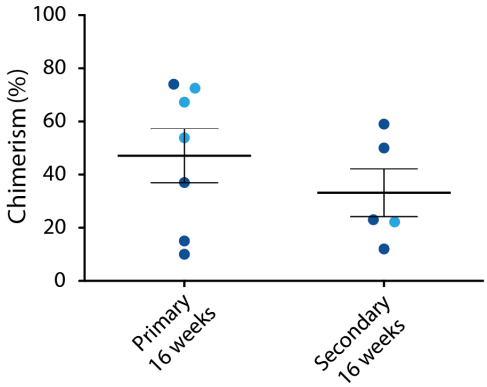


Figure 35 Comparison of 16 week chimerism in primary and secondary recipients of WT bone marrow
Previously presented data (light blue, as in Figure 28), and the additional data (dark blue) are not significantly different in primary transplantation (p=0.141). Error bars are SEM.

significance of JAK HOM TET HET (without extra data p=0.003, with extra data p=0.090). The additional data was therefore included for comparison of chimerism in the secondary transplantation. With the additional data secondary chimerism was 16 weeks 33.2%.

Recipients of secondary TET HOM BM follow the trends established in the primary recipients. Average chimerism at 16 weeks was similar to that of WT recipients (Figure 37 left, 45.54%, p=0.610), however as in the primary, there was a wide variation with 2/5 samples with very high chimerism (over 95%) and 3/5 with quite low chimerism (under 20%). Once again, the haemoglobin and haematocrit were within normal range (Figure 37 middle and right), as loss of TET2 does not result in a hyperproliferation of the erythrocyte lineage (at 16 weeks post-transplantation HGB 13.82g/dL, and HCT 54.6%). As TET2 blood parameters have no erythrocytic phenotype, and blood parameter data was not available for the additional WT data set, TET HOM recipients were used for statistical comparison with JAK2 mutant models.

Secondary recipients of JAK HOM BM continue to have very low chimerism in the secondary transplantation, with an average chimerism of 3.60% at 16 weeks post-transplantation (p=0.030) (Figure 38 left). As in the primary animals, the lack of chimerism resulted in haematocrit and haemoglobin within the normal range as there were not a sufficient number of mutant cells to drive elevation of erythroid cell numbers; at 16 weeks post-transplantation average haematocrit of JAK HOM recipients was 56.8% (p=0.320), and average haemoglobin 13.4 g/dL (p=0.186) (Figure 38 middle and right).

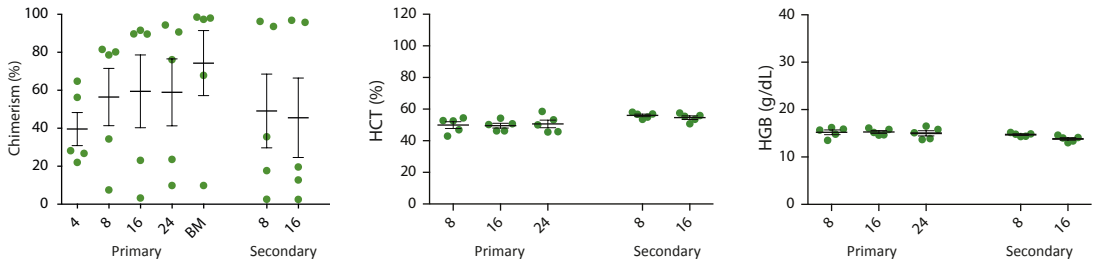


Figure 37 Chimerism and blood parameters of secondary recipients of TET HOM BM in competitively transplant setting
Secondary recipients of TET HOM BM continued to have variable chimerism and HGB and HCT within the normal range. Error bars are SEM.

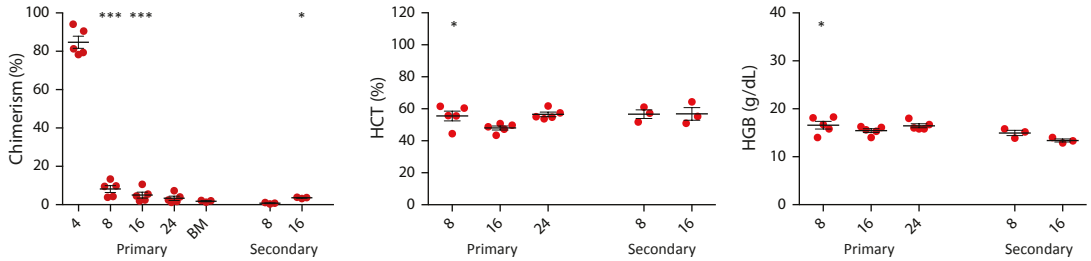


Figure 38 Chimerism and blood parameters of secondary recipients of JAK HOM bone marrow in competitive transplant setting

Secondary recipients of JAK HOM BM continued to exhibit the effects seen in primary recipients; significantly lower chimerism (16 weeks p=0.030) and no changes in HCT or HGB (relative to TET HOM, HCT p=0.320, HGB p=0.186 at 16 weeks). Error bars are SEM.

The chimerism in primary recipients of JAK HOM TET HET underwent a slow decline, much slower than the exhaustion seen in JAK HOM alone but a gradual decrease in chimerism with time. This continued into the secondary recipients and by 16 weeks post-transplantation average chimerism was 5.66% (p=0.036) (Figure 39 left). As in JAK HOM alone the chimerism in the secondary recipients of JAK HOM TET HET was insufficient to sustain elevated haematocrit and haemoglobin out to 16 weeks post-secondary transplantation (Figure 39 middle and right, average HGB 13.275g/dL, p=0.242, and average HCT 50.65%, p=0.176). Indicating that although loss of a single copy of TET is able to slow the exhaustion of JAK HOM HSCs, it is not able to reverse the self-renewal defect entirely and consequently is not capable of imparting a serially transplantable MPN-like phenotype.

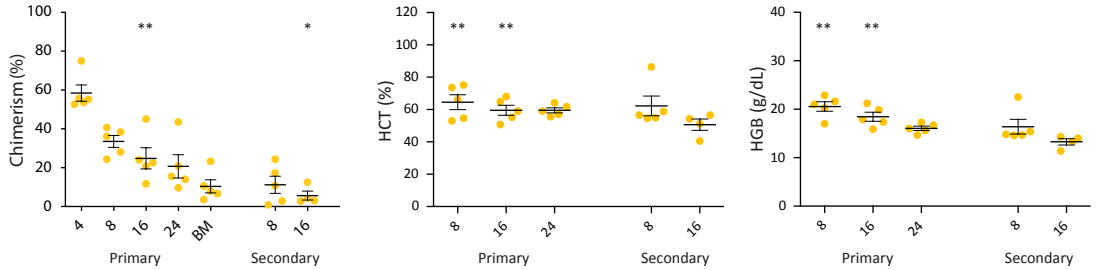


Figure 39 Chimerism and blood parameters of secondary recipients of JAK HOM TET HET bone marrow in competitive transplant setting

Chimerism continued to decline in secondary recipients of JAK HOM TET HET BM, by 16 weeks chimerism was significantly lower than WT (p=0.036), this chimerism was insufficient to cause elevation in HCT (p=0.176) or HGB (p=0.242). Error bars are SEM.

The chimerism of JAK HOM TET HOM recipients resembles that of TET HOM alone, with a similar average chimerism to WT (Figure 40 left, at 16 weeks post-secondary transplantation 32.52%, p=0.974), but a large spread in the data within the chimerism of individual recipients at 16 weeks ranging from 2% to 80%. The primary recipients of JAK HOM TET HOM BM displayed striking increases in haematocrit and haemoglobin when compared with WT recipients; in the secondary transplantation this erythrocytic phenotype continues to be seen in the 8 week bleed (Figure 40 middle and right, average HGB 22.51g/dL, p=0.002, and average HCT 85.35%, p>0.001) and 16 week bleed (average HGB 19.61g/dL, p=0.039, and average HCT 72.20, p=0.038).

At 18 weeks post-secondary transplantation mice were sacrificed and BM analysed by flow cytometry. The BM from the secondarily transplanted recipients summarises the effects of each mutation upon self-renewal. TET2 loss alone recipients had a strong average chimerism with a normal distribution of cell types (Figure 41 left). JAK2 V617F alone has a self-renewal defect, exhausting quickly, and was hence unable to sustain chimerism out to the end of the secondary transplant, by this point all recipients have donor chimerism of less than 1% - commonly accepted as the cut off for positive engraftment. By combining JAK2 V617F with loss of a single copy of TET2 self-renewal improves slightly, with a more prolonged drop in donor chimerism over the course of the experiment, but by the end of the secondary transplantation 2/4 grafts has exhausted and other 2 were very low (1.13 and 3.44%). As in primary transplantation for JAK HOM the balance of cell types within these grafts are more weighted towards B and T lymphocytes (Figure 41 middle), this is not a lymphoid bias but a consequence

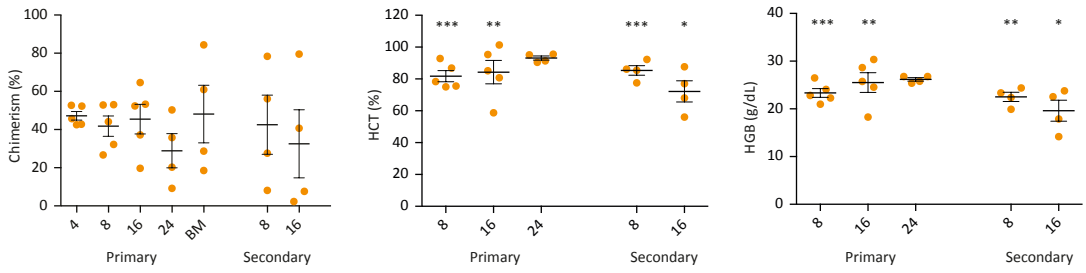


Figure 40 Chimerism and blood parameters of secondary recipients of JAK HOM TET HOM bone marrow in competitive transplant setting

Donor chimerism reveals JAK HOM TET HOM grafts have not exhausted by 16 weeks post-secondary transplantation. Elevated HCT and HGB continue to be seen in secondary recipients of JAK HOM TET HOM BM (at 16 weeks post-transplantation, HCT p=0.038, HGB p=0.039). Error bars are SEM.

of these cells having a longer life and therefore contributing to chimerism for longer. The loss of both copies of TET2 improves the self-renewal of JAK HOM mutant HSCs further, mice are able to sustain chimerism out to the end of the secondary transplantation, and these grafts contain the highest average proportion of phenotypic HSCs (although this was not significantly higher than other models) (Figure 41 right).

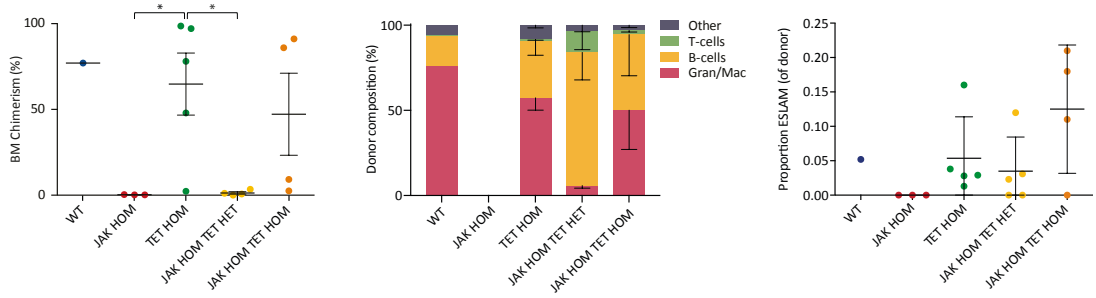


Figure 41 Assessment of the bone marrow chimerism and composition at the end of competitive secondary transplantation

As in primary transplantation BM, chimerism was significantly lower in JAK HOM (p=0.023) and JAK HOM TET HET (p=0.025) than in TET HOM, but was not significantly lower in JAK HOM TET HOM than in TET HOM (p=0.577). No significant differences were observed in the proportion of different lineages within donor grafts, or the proportion of ESLAM cells within each graft. Error bars are SEM.

4.4 JAK2 mutation reduces the proportion of ESLAM HSCs in steady state bone marrow

To assess whether the restoration of the JAK2 V617F self-renewal defect, by combining with loss of TET2, impacts the number of phenotypic HSCs, the proportion of ESLAM and Sca1^{+} c-Kit^{+} (SK) progenitor cells from lineage depleted BM was calculated relative to WT control experiments (relative number, rather than absolute number, is reported to control for difference in lineage depletion/staining etc between experimental days). In agreement with previous studies[228], JAK HOM animals had a reduced number of phenotypic HSCs (ESLAM) and an increase in the proportion of progenitor cells (SK) (Figure 42). Hyperproliferation of progenitors leads to increase in the progenitor population, and the increase in the number of progenitors and downstream mature cells causes a proportional reduction in number of phenotypic HSCs.

TET mutant animals had a similar (slightly higher) proportion of ESLAM HSCs and SK HSPCs to WT animals (Figure 42). This is in agreement with previously published data on this mouse model; Ko et al. (2011)[210] reported an increase in the absolute number of LK and LSK cells in TET2 knock-out animals.

Combination JAK and TET mutant animals, like JAK2 mutation alone, had a reduced frequency of ESLAM HSCs (TET HOM vs JAK HOM TET HET p=0.045), this is presumably due to the increased proportion of mature progeny invoked by this hyperproliferating mutation. Double mutant mice contained a decreased frequency of SK cells, differing from JAK2 mutation alone (JAK HOM v JAK HOM TET HET p=0.029).

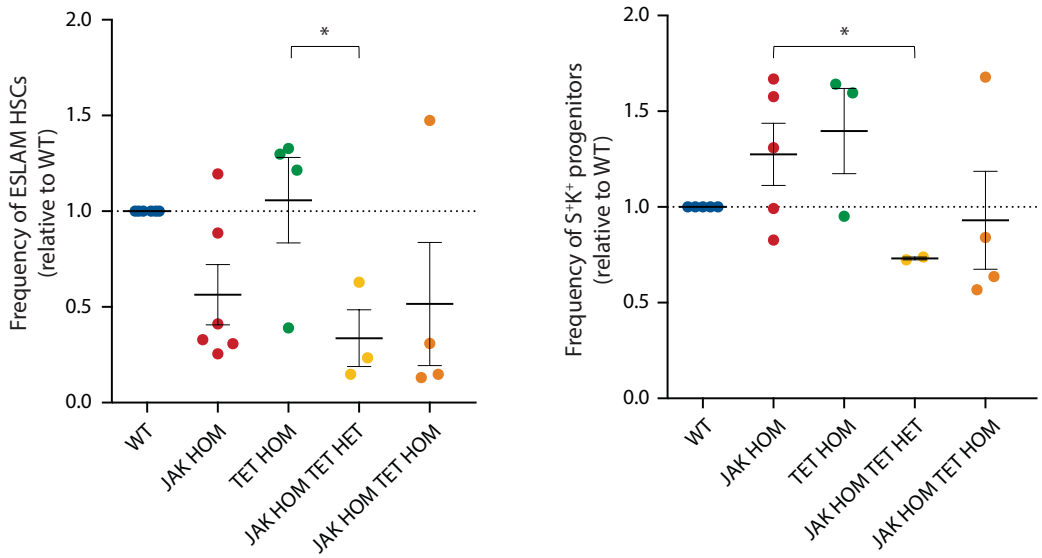


Figure 42 Relative number of phenotypic stem and progenitor cells in mice with JAK2/TET2 mutation

To mitigate the differences in lineage depletion/staining etc between experiments, values are relative to WT mice processed at the same time. JAK HOM TET HET mice had proportionally fewer ESLAM HSCs than TET HOM mice (p=0.045), and proportionally fewer SK progenitors that JAK HOM mice (p=0.029). Error bars are SEM.

Results

5

Molecular mechanisms driving self-renewal in MPN HSCs

In this chapter I explored the molecular mechanisms driving the self-renewal defect of JAK2 V617F HSCs, by using single cell multiplexed qPCR to examine the gene expression of 45 self-renewal regulators to attempt to identify individual genes driving the defect. Identified genes were then validated by over-expression using lentiviral vectors to determine whether these genes could restore the self-renewal defect. The expression of these genes was also investigated in bulk mouse cells post-transplantation and MPN patient samples. In section 5.1 the multiplexed qPCR was performed by Jiangbing Li and David Kent, and analysis performed by Nicola Wilson

5.1 Single- and double-mutant HSCs have distinct molecular profiles of self-renewal regulators

In order to understand the mechanisms by which a malignant clone can expand and drive disease, an appreciation needs to be gained of how malignant HSCs differs from normal HSCs; which molecular programmes they have employed to evade normal regulation and balance. Study of the genes expressed within mutant cells gives an insight into how mutations drive different functional characteristics and can provide potential targets for disease treatments.

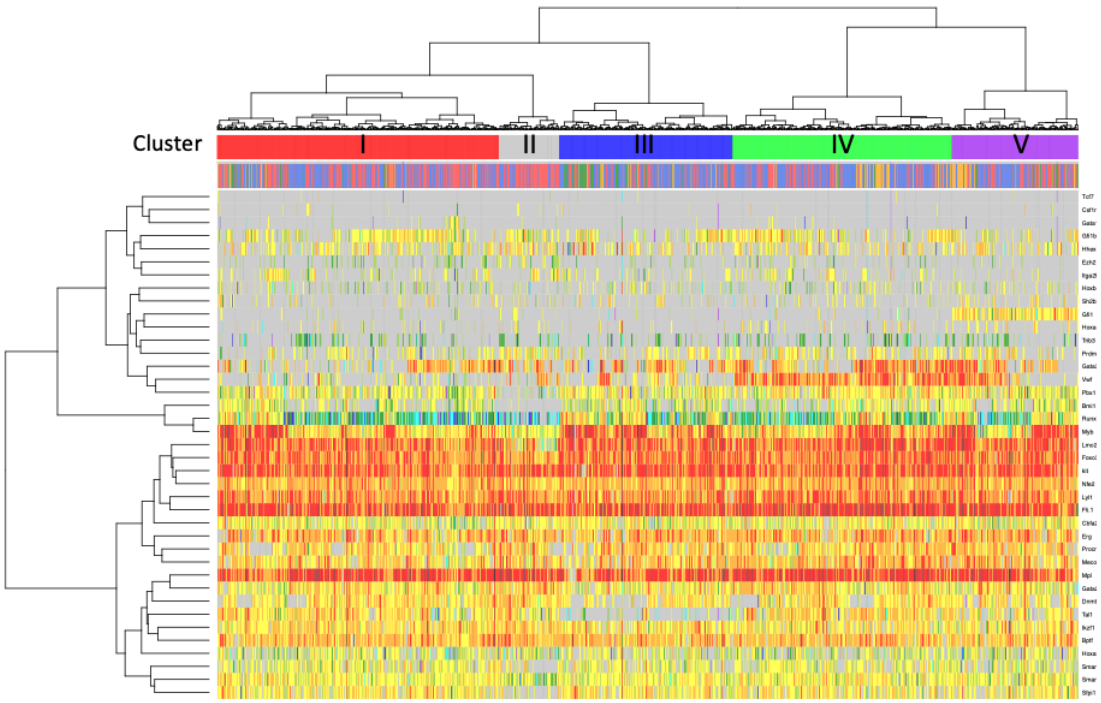
5.1.1 Hierarchical clustering identifies subtypes within HSC population that are enriched in different genotypes

The transplantation data in chapter 4 identifies a self-renewal defect in in JAK2 mutant HSCs, this defect is rescued by homozygous deletion of TET2. To determine whether expression of self-renewal genes is altered by these mutations, individually or in combination, single phenotypic HSCs from WT (N=3, n = 561), JAK HOM (N=3, n = 376), TET HOM (N=1, n = 73) and JAK HOM TET HOM (N=1, n = 99) mice were analysed for the expression of 45 transcription factors and self-renewal regulators by multiplexed quantitative polymerase chain reaction. These 45 genes were selected from literature and are listed, and their roles in self-renewal discussed, in Table 1. Quantitative analysis of 48 genes (3 housekeeping genes plus 45 self-renewal implicated genes) was performed using Fluidigm 48.48 Dynamic Array chips, this allowed quantitative analysis of all genes in parallel at the single-cell level. For this, single ESLAM HSCs were sorted, followed by lysis, cDNA synthesis, and sequence-specific preamplification all within the sorted well. Quantitation of gene expression was performed using TaqMan real-time PCR on the Fluidigm BioMark system.

Hierarchical clustering shows 5 broad molecular clusters of cells (Figure 43). Clusters I and II are significantly enriched ($p< 0.05$ and $p< 0.005$) for JAK HOM HSCs, Cluster III is significantly enriched ($p< 0.001$) for TET HOM HSCs, and Cluster IV is significantly enriched for WT ($p< 0.05$) and double-mutant cells ($p< 0.05$). Interestingly, Cluster II does not have any TET single- or double-mutant HSCs. WT HSCs are present across all clusters; this exemplifies the heterogeneity of the normal HSC population. This spread likely represents a combination of informative and generic heterogeneity. As gene expression is measured at a snapshot in time, a degree of the differences between HSCs may correspond to the differences in cell state, like stage of cell cycle (generic heterogeneity), however as the genotypes with their different characteristics are enriched in distinct clusters the difference in their gene expression is most likely biologically driven, and the presence of WT cells in each cluster represents informative heterogeneity.

Even from a visual inspection of the hierarchical clustering, some of the genes defining each cluster are clear, for example expression of *Gfi1* is almost exclusively found in cluster V, lack of *Tal1* expression is more common among cluster III cells, cluster IV

cells have on average higher expression of *Vwf*, and cluster II cells largely lack expression of *Bmi1* (Figure 43). This sets the stage for assessment of which genes are associated with each individual genotype by multidimensionality reduction.



	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
WT	0.4804	0.1252	0.7124	0.04	0.2938
JAK HOM	0.0282	0.0028	0.0333	0.0057	0.988
TET HOM	0.4677	0.0687	0.0005	0.0015	0.6604
DOUBLE	0.0512	0.0209	0.4562	0.0482	0.0795

Figure 43 Hierarchical clustering of single cell gene expression data
Five main clusters emerge when single HSCs from mice with each of the four genotypes (WT, JAK HOM, TET HOM, and JAK HOM TET HOM) are assessed. Individual HSCs and their expression for each self-renewal regulator with HSCs from each genotype are coloured accordingly: Blue (WT), Red (JAK HOM), Green (TET HOM) and Orange (double mutant). Each cluster was assessed for whether or not it contained an increased or decreased number of HSCs from each genotype compared to that expected by random chance by chi-squared analyses and these results are displayed in the table (yellow indicates $p < 0.05$ of a higher than expected observation and purple indicates $p < 0.05$ a lower than expected observation).

5.1.2 Principal component analysis reveals genes associated with each genotype

Single HSCs of all genotypes were displayed on a single PCA plot to help visualise molecular clusters and identify the genes associated with each genotype (Figure 44). WT HSCs existed in all regions of the molecular landscape (Figure 44 top left), whereas JAK HOM HSCs were overrepresented in a specific region (Figure 44, top right), and this region was almost completely devoid of cells lacking TET2 expression, suggesting it is a unique molecular region exclusive to a minority of WT HSCs and a larger proportion of JAK HOM HSCs. The majority of these cells are also found in Cluster II described above. The single-mutant TET HOM HSCs (Figure 44 middle left) are predominantly found in a single area of the PCA plot, this region is enriched for cluster III cells. The double-mutant HSCs are enriched in another region (Figure 44, middle right), although a small fraction of double-mutant cells are also found with the single-mutant TET HOM HSCs. The genes negatively associated with each of these regions are indicated in the loadings plot (Figure 44 bottom left), implicating genes such as *Bmi1*, *Runx1*, and *Pbx1* as being involved in the self-renewal of HSCs (ie, associated with TET HOM and double-mutant genotypes), rather than proliferation (ie, associated with JAK HOM HSCs).

When t-Distributed Stochastic Neighbour Embedding (t-SNE) analysis is used the same patterns can be seen as with PCA. WT cells occupy the entire molecular landscape (Figure 45 top left), JAK HOM cells are overrepresented in a region (Figure 45 top right), in which TET mutant cells are completely absent. TET HOM cells cluster in a separate region (Figure 45 bottom left), and JAK HOM TET HOM cluster in another region (Figure 45 bottom right) but are also found in the TET HOM enriched region. The same pattern of genotype separation in both the PCA and t-SNE gives confidence in the differences between HSCs of each genotype and the genes driving these differences.

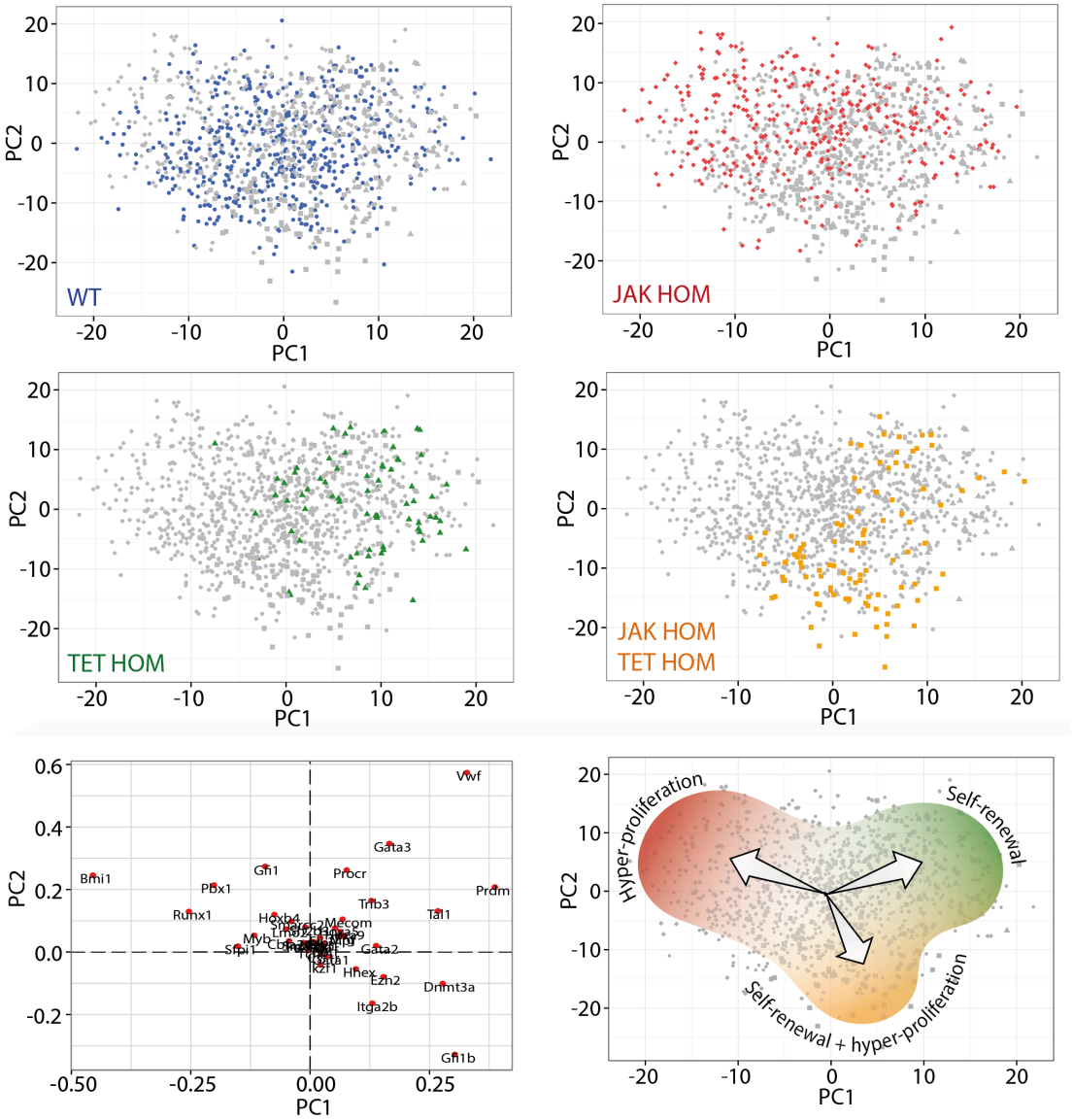


Figure 44 PCA of single-cell gene expression profiling shows distinct molecular regions of single and double-mutant HSCs

Cells from all 4 genotypes were assessed by single-cell multiplexed quantitative polymerase chain reaction; each PCA plot displays all single cells analysed with 1 population highlighted in each plot; HSCs from WT (n=561, blue circles), JAK HOM (n=376, red circles), TET HOM (n=73, green circles), and JAK HOM TET HOM (n=99, orange circles). Notably, WT cells are present across the entire molecular landscape, whereas single- and double-mutant HSCs are enriched in specific regions. Loadings plot for PCA, indicating the key defining molecular features of each region, with the distance from the centre indicating the negative correlation with a cell type (e.g., cells in the top right lack Vwf). Illustration depicting the different cell characteristics associated with each region of the molecular landscape. Axes are in arbitrary units.

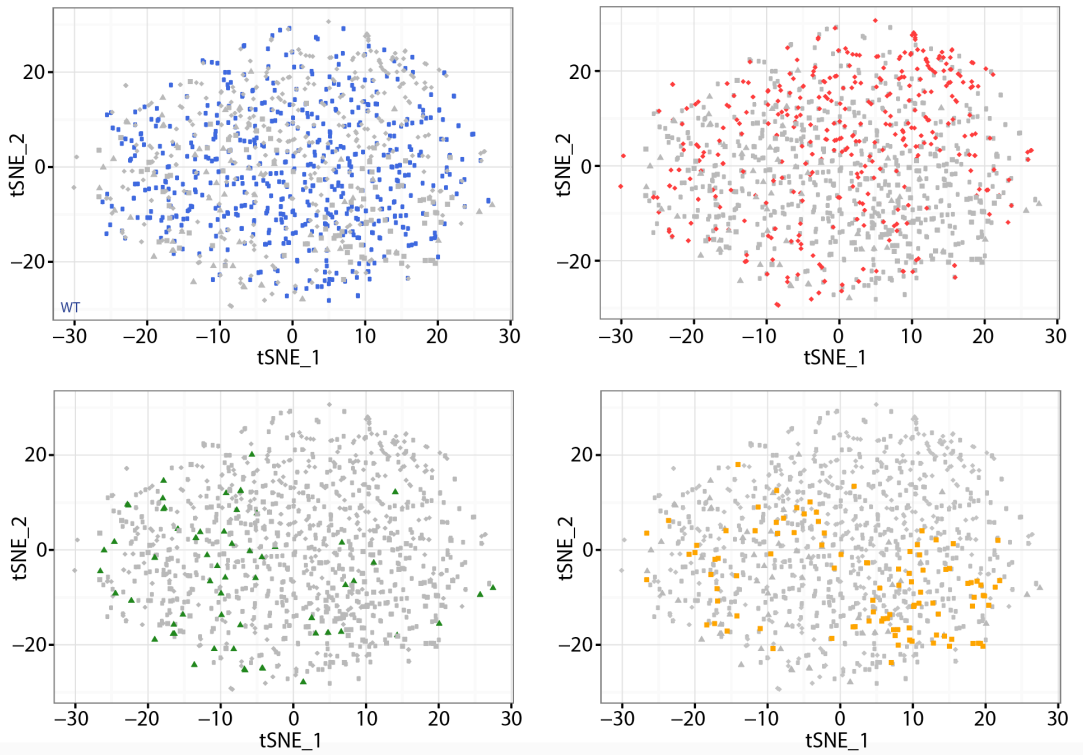


Figure 45 T-SNE shows distinct molecular regions of single and double-mutant HSCs

T-SNE analysis for the same dataset that appears in the PCA in Figure 44 Each t-SNE plot displays all single cells analysed with 1 population highlighted in each plot. T-SNE displays on a single plot, HSCs from WT (blue squares, n=561), JAK HOM (red diamonds, n=376), TET HOM (green triangles, n=73), and JAK HOM TET HOM (orange squares, n=99) mice.

5.1.3 Identification of candidate genes to 'rebalance' HSC population in JAK HOM HSCs

As the gene expression was performed on single HSCs, the clustering diagram and differential enrichment by genotype, in combination with the PCA loadings plot, can be used to ascertain which genes might define the molecular subtypes, and identify candidates for 'rebalancing' the HSC pool. For example, double-mutant cells (which have restored self-renewal but still proliferate rapidly) cluster away from JAK HOM single-mutant cells (which have high proliferation but lack durable self-renewal), allowing the uncoupling of individual characteristics, the genes driving this separation are only involved in self-renewal, not hyperproliferation. This permits the generation a list of candidate genes that might drive the self-renewal restoration observed in *in vivo* experiments; increased *Bmi1, Pbx1, Runx1*, and *Meis1* and decreased *Gfi1b* and *Dnmt3a*.

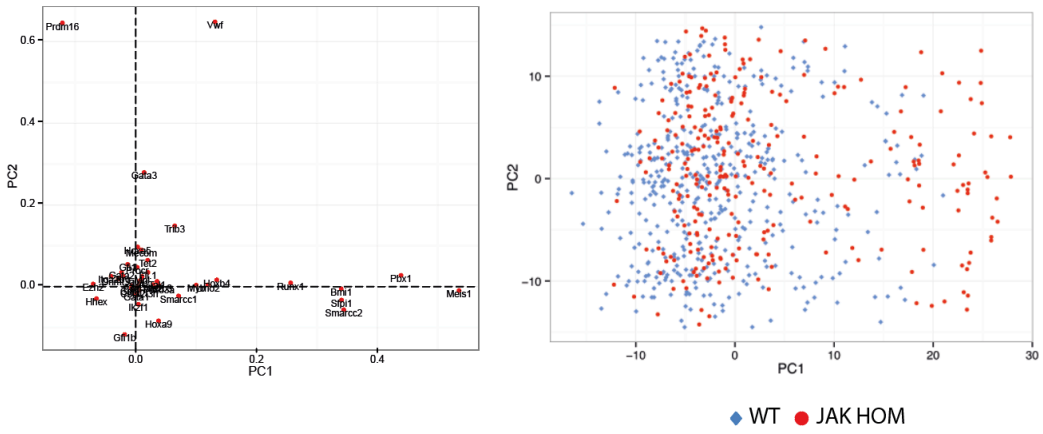


Figure 46 PCA of WT and JAK HOM HSCs

JAK HOM HSCs are indicated by red circles, and WT HSCs are indicated by blue diamonds. A cluster of cells on the right hand side of the graph is enriched for JAK2 V617F HOM HSCs and have reduced expression of several key HSC genes (Meis1, Pbx1, Bmi1, Sfpi1, Smarcc2, Runx1, Hoxb4, Myb, Lmo2). Axes are in arbitrary units. WT, n=465; JAK HOM, n=277.

It should be noted that *Meis1* is included in this list despite being absent from the PCA loadings plot. This is because *Meis1* was excluded from the PCA having failed the loading control (in the run with TET HOM and JAK HOM TET HOM), however *Meis1* was the strongest driver of 'abnormal' JAK2 cells in runs of JAK HOM and WT alone (Figure 46) and was therefore included for subsequent validation.

The individual expression patterns of the identified genes, hypothesised to be driving self-renewal, across genotypes are displayed by violin plots (Figure 47). *Bmi1* expression is bimodal for all genotypes, but the proportion of cells in each group varies between genotypes; for JAK HOM and WT a larger proportion of cells have low/no expression of *Bmi1* than for TET HOM and JAK HOM TET HOM.

Like *Bmi1*, *Pbx1* expression is bimodal for all genotypes and JAK HOM and WT have a larger population of cells with low *Pbx1* expression than TET HOM and JAK HOM TET HOM.

All TET HOM and the vast majority of JAK HOM TET HOM cells have high expression of *Runx1*. WT and JAK HOM cells have some cells with high *Runx1* expression but a larger population of cells with medium expression and another population of cells with low expression. *Meis1* is expressed in all TET HOM and JAK HOM TET HOM cells but, in a proportion of JAK HOM and WT cells, *Meis1*

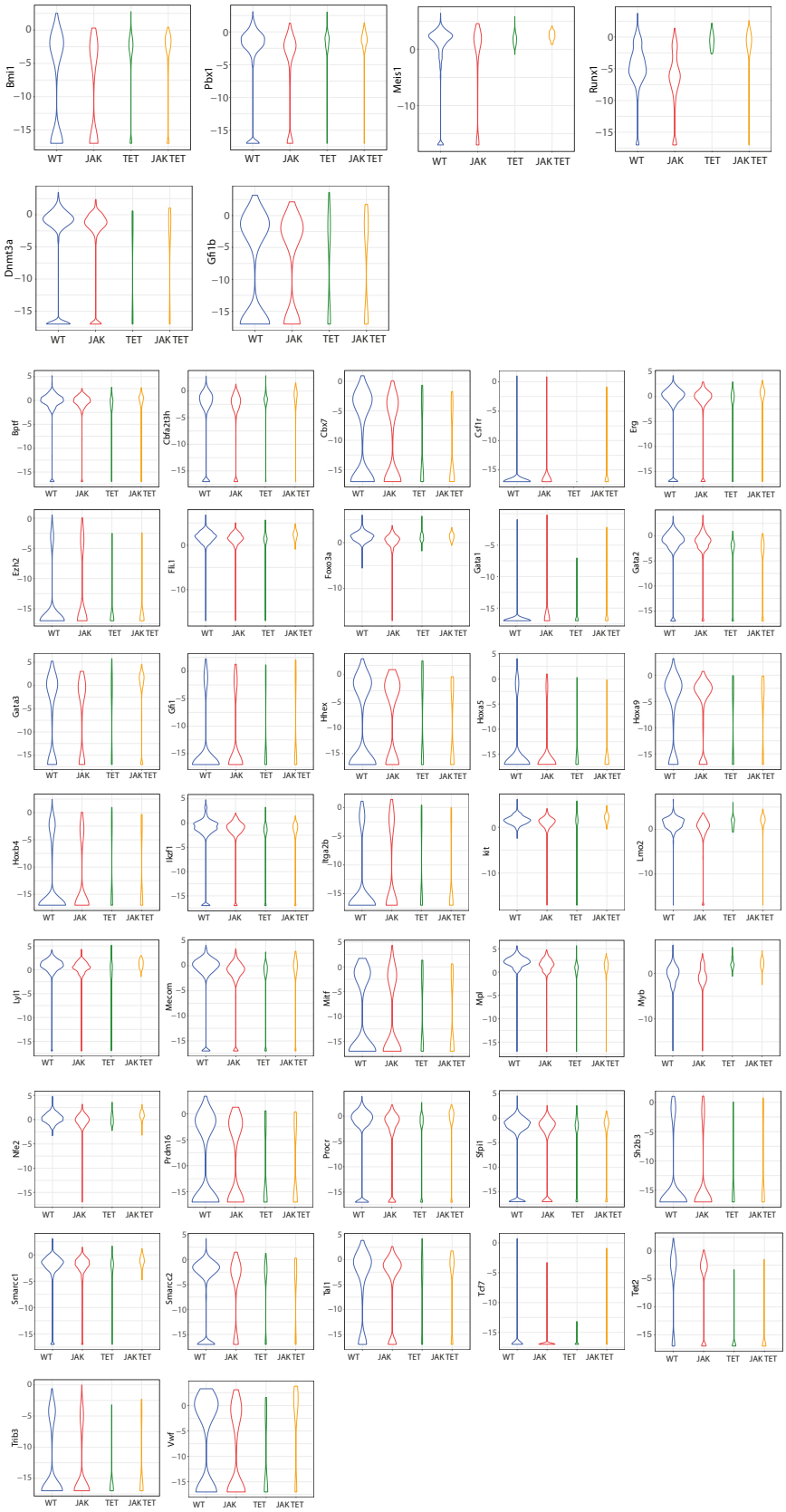


Figure 47 Violin plots for expression of each of the 45 self-renewal regulators in each genotype

Top panel shows genes differentially regulated in JAK HOM HSCs; Bmi1, Pbx1, Meis1, Runx1, Dnmt3a, and Gfi1b, all others genes in bottom panel.

expression is very low/absent. However, Meis1 did fail its loading control for TET HOM and JAK HOM TET HOM so it is unknown if the presence of expression in these cells is real and no conclusion can be drawn from this.

The genes for which decreased expression correlates with a self-renewal advantage, *Dnmt3a* and *Gfi1b*, both have bimodal expression patterns with a larger proportion of WT and JAK HOM HSCs having high expression of these genes than TET HOM and JAK HOM TET HOM HSCs. Between TET HOM and JAK HOM TET HOM HSCs *Vwf, Gata3, Prdm16,* and *Tal1* have reduced expression in TET HOM HSCs (but not reduced expression in JAK HOM TET HOM)

5.2 Candidate gene validation

A reminder from above – the genes that drive the self-renewal restoration of JAK2 V617F self-renewal defect; increased *Bmi1, Pbx1, Runx1*, and *Meis1* and decreased *Gfi1b* and *Dnmt3a*. As loss of DNMT3A is a frequent mutation in MPNs, ARCH and other haematological malignancies, and frequently commutated with JAK2 V617F in MPNs, knock-down/out of DNMT3A does not seem like an appropriate means of validating the mode of action of TET2 mutation. This caused me to focus on the genes that are under-expressed in JAK HOM HSCs; *Bmi1, Pbx1, Runx1*, and *Meis1,* for which over-expression could restore the self-renewal defect of JAK HOM HSCs.

Genes of interest were engineered into a pCCL-c-MNDUS-(X2)-PGK-EGFP backbone. The gene of interest was under the MND promoter (myeloproliferative sarcoma virus with the negative control region deleted), this is constitutively active in the haematopoietic system and commonly used for generating overexpression in this tissue; in a study to identify the most potent promoter in the haematopoietic system for gene editing for Wiskott–Aldrich syndrome, the MND promoter was found to be more effective than EF1α and WASp, driving higher expression in all lineages.[264] The vector also contained eGFP, under the control of PGK promoter, this acted as a selection marker allowing transduced cells to be identified.

5.2.1 Overexpression of *Bmi1, Pbx1*, or *Meis1* can help sustain JAK HOM chimerism *in vivo*

To test which of the self-renewal genes identified as being under-expressed in JAK HOM HSCs (*Bmi1, Pbx1, Runx1*, and *Meis1*) might modulate the self-renewal of these cells, lentiviral overexpression and transplantation assays were performed. $CD45^{+}Lin^{-}CD150^{+}CD48^{-}$ HSPCs isolated from either JAK HOM or WT mice were transduced with a lentivirus containing a GFP reporter and the *Bmi1*, *Pbx1*, *Runx1*, or *Meis1* genes or an empty vector (EV) control. Three days post-infection, cells were re-sorted for presence of GFP and transplanted into recipient mice to monitor for disease phenotype (Figure 48).

For transplantation of cells over-expressing genes, the reporting of chimerism and blood parameters post-transplantation is slightly more complicated than for previous whole BM transplantation. This is because within each group there are mice which have not successfully repopulated (chimerism below 1%), the lack of repopulation could be the result of a number of factors. Firstly, the sorted population was transduced with lentivirus contained both stem and progenitor cells, it is not known which of these cells are successfully transduced and transplanted, should the population be mostly/entirely progenitors the graft may not able to be sustained. Secondly, the cells were kept in culture for three days before transplantation, during this time the cells divide, and to varying degrees differentiate. Even if many HSCs are successfully transduced, the

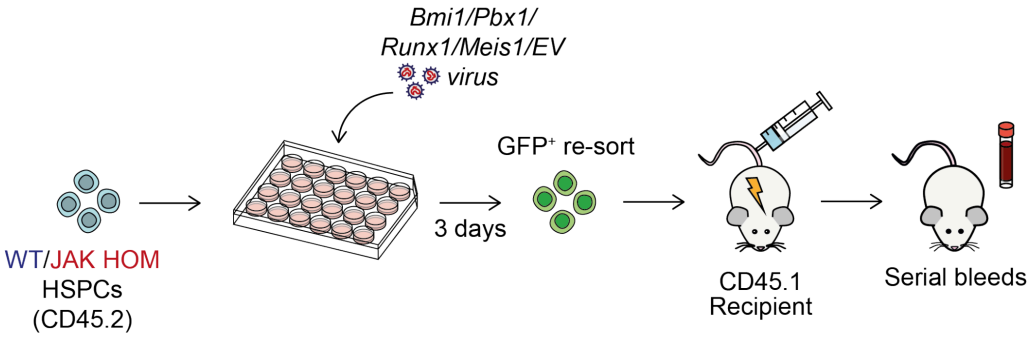


Figure 48 Schematic of candidate gene overexpression transplants
Bulk $CD45^{+}Lin^{-}CD48^{-}CD150^{+}$ (HSPC) cells were sorted from WT and JAK HOM mice and infected with lentivirus carrying no gene (empty vector) or lentivirus carrying genes to overexpress Bmi1, Pbx1, Runx1, or Meis1 (2 independent experiments). Three days after infection, GFP^{+} cells (300-2000) were isolated and transplanted into recipient mice. Serial analysis of peripheral blood was performed to assess chimerism and blood cell parameters.

culture conditions could drive them to differentiate and by transplantation the population may be too differentiated to sustain engraftment. Thirdly, the stress of the viral transduction can cause cells to die/differentiate. This combined with low cell numbers and a mixed starting population could mean no HSCs are left by transplantation injection. The repopulation success is still a measure of a given gene's effect upon cells, however it also misleads the average chimerism by pulling down the average of successfully transplanted mice. Both the average chimerism, repopulation success rate, and chimerism of individual mice will be reported to try to give a more complete picture of the effect of each gene on WT and JAK HOM HSPCs.

WT and JAK HOM HSPCs infected with an empty vector lentivirus had comparable repopulation success (WT 4/5, JAK 3/4 +1 dead) and, at early time points, similar average chimerism (8 weeks, Figure 49, WT average 24.85%, JAK2 average 20.00% p=0.747). This remained true out to 16 weeks post-transplantation, after this time JAK HOM EV average chimerism began to fall and by 24 weeks JAK HOM EV average chimerism was 6.00%, whereas WT was 23.65% (p=0.186)(Figure 49).

This donor graft exhaustion parallels the effects seen in JAK HOM whole BM transplantation (section 4.3) but in this case exhaustion begins later (16 weeks post transplantation rather than 4 weeks). This difference in time to chimerism peak is likely due to the difference in the transplanted cell population. In this case HSPCs were

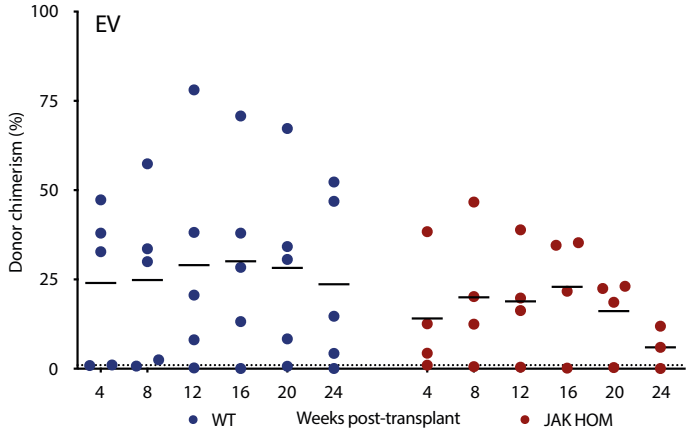


Figure 49 Peripheral blood chimerism of recipients of WT and JAK HOM cells infected with EV lentivirus

WT in blue, JAK HOM in red. Although chimerism between individual animals was variable, average chimerism in WT mice remained relatively stable throughout the experiment while JAK HOM chimerism began to fall after 16 weeks.

transplanted after 3 days in culture, this donor population differs in two important ways from whole BM. Firstly, this population most likely contains more stem and progenitor than $5x10^5$ whole BM (long-term HSC frequency in normal WT WBM is 1 in 20 000, in 500 000 WBM cells would expect 25 HSCs). Secondly, this more primitive population, and smaller overall number of mutant cells transplanted, may differently impact the BM microenvironment. The large number of differentiated cells within JAK HOM BM will create a specific microenvironment, which could encourage the proliferation and differentiation of stem and progenitor cells thereby impacting self-renewal.

Overexpression of *Runx1* resulted in a much lower rate of successful transplantation in recipients of both WT and JAK HOM cells (at 16 weeks post-transplantation 1/5 had chimerism over 1% for both WT and JAK HOM), and the successfully engrafted mice had very low chimerism, at 24 weeks post-transplantation average chimerism for WT *Runx1* recipients was 2%, (WT EV 23.65%, p=0.118), and JAK HOM *Runx1* recipients 0.38% (JAK HOM EV 6.00%, p=0.240) (Figure 50).

Recipients of *Pbx1* overexpression cells also had a lower proportion of successful transplantation than EV recipients, WT 1/5 and JAK HOM 2/5. Of the successfully engrafted recipients, the WT recipient had very strong chimerism across all timepoints (Figure 51, at 16 weeks 60.50%), of the JAK HOM positive transplantation recipients

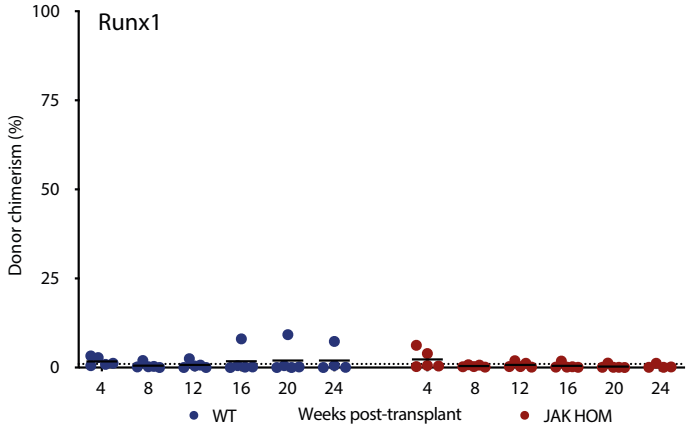


Figure 50 Peripheral blood chimerism of recipients of WT and JAK HOM cells overexpressing Runx1
Recipients of WT cells in blue and JAK HOM in red. Low chimerism for all recipients of WT and JAK HOM cells overexpressing Runx1 (although due to variability of EV not significant, compared equivalent EV at 24 weeks, WT Runx1 p=0.118, JAK HOM Runx1 p=0.240).

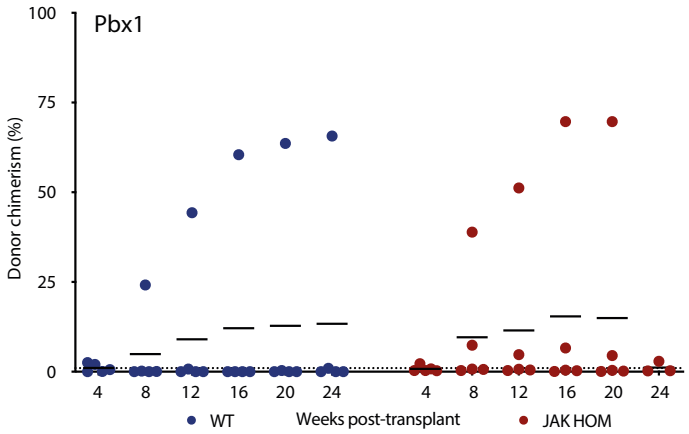


Figure 51 Peripheral blood chimerism of recipients of WT and JAK HOM cells overexpressing Pbx1
Recipients of WT cells in blue and JAK HOM in red. For both WT and JAK HOM, one mouse was able to generate high donor chimerism, while donor chimerism in the other recipients was low. Compared equivalent EV at 24 weeks, WT Pbx1 p=0.562, JAK HOM Pbx1 p=0.287.

one had continually low chimerism (at 16 weeks 6.63%), while the other had very high chimerism up to 16 weeks (chimerism 69.70%) when this animal had to be culled. The mouse displayed partial paralysis/jerky movements, a blood sample taken before death revealed extremely high HCT and HGB (HCT 96.02%, HGB 24.20g/dL), and it was therefore hypothesised that this mouse suffered a stroke due to the severely elevated erythrocytosis (unconfirmed as no analysis of brain performed).

The varying effects of *Pbx1* overexpression, in some instances not leading to any engraftment and in others very strong engraftment, could be due to a mixed starting population of stem and progenitor cells; for instance, *Pbx1* expression in progenitors or primed/biased HSCs might results in differentiation (and thus reduced positive engraftment) while the same expression in a more primitive HSCs or an HSC with a different bias/priming might give an advantage, and when combined with JAK2 V617F this is able to give a myeloproliferative phenotype.

All (5/5) recipients of WT cells overexpressing *Meis1* were successfully repopulated and 4/5 recipients of *Meis1* JAK HOM had chimerism over 1%, equivalent repopulation to EV (Figure 52). In WT Meis1 recipients average chimerism was similar to WT EV across all timepoints (at 16 weeks p=0.909). Compared to EV JAK HOM, *Meis1* JAK HOM chimerism was lower at earlier time points (at 4 weeks EV 14.09%, Meis1

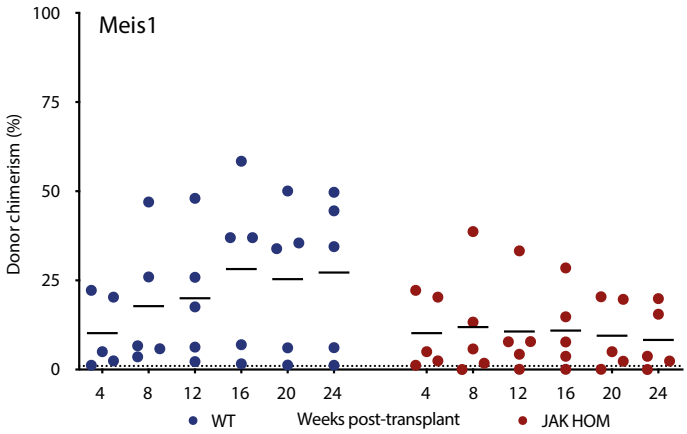


Figure 52 Peripheral blood chimerism of recipients of WT and JAK HOM cells overexpressing Meis1
Recipients of WT cells in blue and JAK HOM in red. Both WT and JAK HOM recipients had chimerism similar to EV controls, compared equivalent EV at 24 weeks, WT Meis1 p=0.815, JAK HOM Meis1 p=0.426.

13.06%, p=0.923), but unlike JAK HOM EV where chimerism begins to fall after 16 weeks post-transplantation, the chimerism in recipients of *Meis1* JAK HOM remained relatively stable, although as this level was lower to begin with this is not significant. Unlike *Runx1* and *Pbx1* which, in many cases, seems to induce differentiation leading to reduced repopulation, over-expression of *Meis1* does not appear to have a negative effect upon repopulation, and may have a slight positive effect, not resulting in increased chimerism but potentially dampening the fall in chimerism seen in JAK HOM alone.

Overexpression of *Bmi1* in WT cells resulted in successful engraftment in 3/5 recipients, average chimerism was lower than for equivalent EV, although chimerism seemed to be climbing throughout time course in recipients of WT *Bmi1* (Figure 53). All recipients of JAK HOM cells overexpressing *Bmi1* were positively engrafted. Like JAK HOM EV, chimerism of these animals grew throughout early time points, after which the chimerism for most mice (3/4) begins to fall – although this is potentially at a slower rate than equivalent EV recipients – unlike EV the chimerism of one mouse continued to climb in the later timepoints, suggesting that under certain conditions *Bmi1* is capable of restoring JAK HOM self-renewal defect. One JAK HOM *Bmi1* recipient was culled after 12 weeks, this mouse displayed the same jerky movement/paralysis as the JAK HOM *Pbx1* recipient and is also suspected of having suffered a stroke.

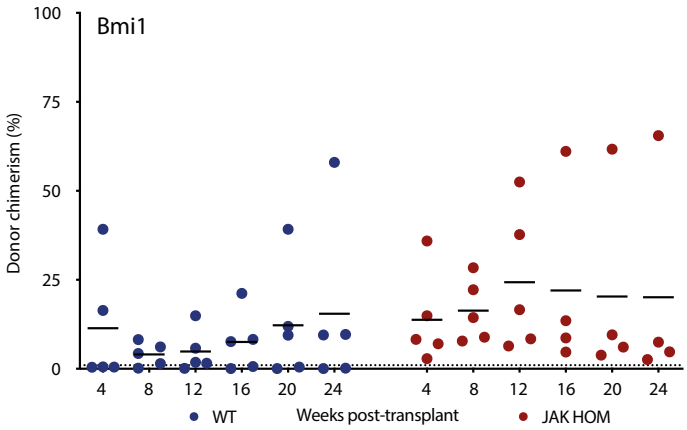


Figure 53 Peripheral blood chimerism of recipients of WT and JAK HOM cells overexpressing Bmi1

Recipients of WT cells in blue and JAK HOM in red. Bmi1 overexpression resulted in positive engraftment in all JAK HOM recipients. Chimerism compared with equivalent EV at 24 weeks, WT Bmi1 p=0.610, JAK HOM Bmi1 p=0.426.

5.2.2 Sustained chimerism leads to MPN phenotype in JAK HOM HSCs overexpressing candidate genes *in vivo*

As discussed earlier (limiting dilution section 4.2), in JAK HOM BM transplantation elevated HCT and HGB are reliant on elevated chimerism. This continues to stand true in the setting of HSPCs over-expressing self-renewal genes. In recipients of WT cells there is no myeloproliferative advantage and consequently no relationship between chimerism and blood parameters (Figure 54 top), the line of best fit not significantly deviated from the mean, R-squared of 0.28%, (0% indicates doesn't explain any of the variation around the mean) (HGB p=0.821, Y = -0.003625*X + 14.69, Rsq 0.002282. HCT p=0.800, Y = 0.01807*X + 57.85, Rsq 0.002842). In recipients of JAK HOM cells elevated chimerism is associated with elevated HCT and HGB (Figure 54 bottom), linear regression shows a significant correlative relationship between chimerism and HCT/HGB (HGB p<0.0001, Y = 0.1188*X + 14.79, HCT p<0.0001, Y = 0.4703*X + 60.04), with a R-squared values of 57.00%(HGB) and 51.04%(HCT), indicating that over 50% of the spread in the data set can be explained by the relationship between chimerism and HCT/HGB.

This correlation combined with the individual gene effects on chimerism can be summarised with the haematocrit and haemoglobin at 20 weeks post transplantation (Figure 55). All WT recipients had average haematocrit and haemoglobin very similar

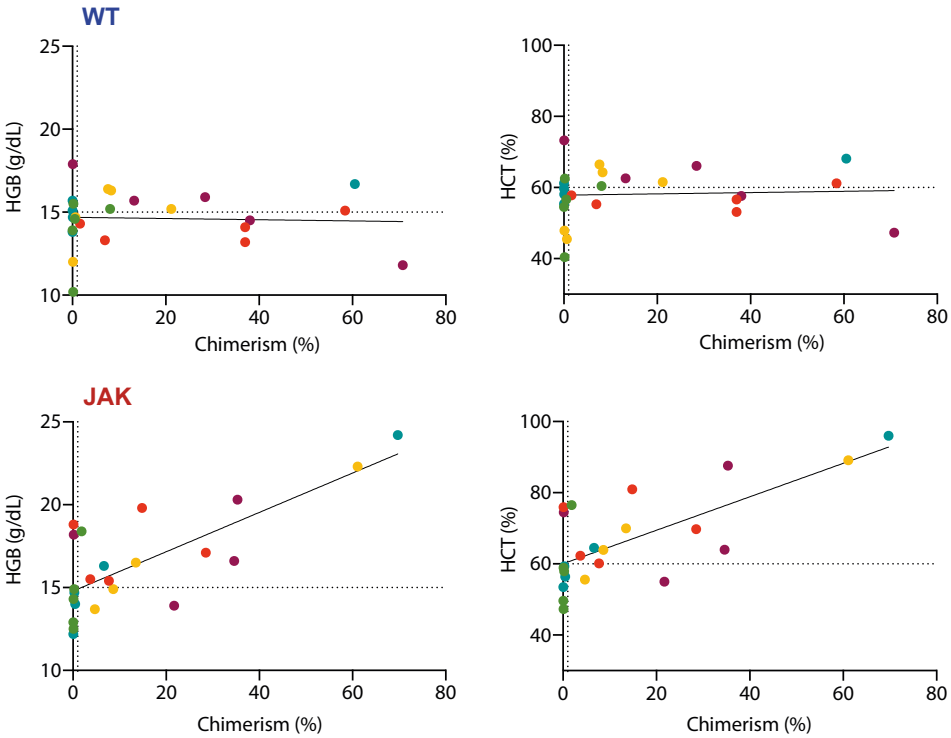


Figure 54 Correlation between chimerism and blood parameters in recipients of WT and JAK HOM cells

In WT recipients (top) no causative relationship between chimerism and HGB (p=0.8206) or HCT (p=0.8002). In JAK HOM recipients (bottom), there is a significant correlative relationship between chimerism and HCT/HGB (HGB p<0.0001, HCT p=0.0001), R-squared values of 57%(HGB) and 51.04%(HCT). Solid line is line of best fit, horizonal dotted line shows average HCT/HGB for WT EV recipients, vertical dotted line at 1% represents the cut off for positive repopulation. Colours correspond with gene being overexpressed; EV=purple, Runx1=green, Meis1=red, Pbx1=blue and Bmi1=yellow.

to WT EV, regardless of repopulation status. JAK HOM EV had average HCT and HGB similar to WT. Likewise, no elevation in HCT or HGB was observed with overexpression of *Runx1*. Mirroring the effects on chimerism, the majority of *Pbx1* overexpressing have no elevation on HCT and HGB but the one mouse with raised chimerism was able to drive a myeloproliferative phenotype (this is the mouse that died of suspected stroke). All positively repopulated JAK HOM *Meis1* animals had above average HCT and HGB; a comparison of repopulated animals showed significant elevation in JAK HOM *Meis1* recipients compared to JAK HOM EV recipients (HGB p=0.049, HCT p=0.019). Overexpression of *Bmi1* in JAK HOM cells, like *Meis1*, resulted in HCT and HGB above average in all recipients. Comparison of positively

repopulated JAK HOM EV mice with JAK HOM *Bmi1* mice revealed this increase was significant (HGB p=0.050, HCT p=0.017).

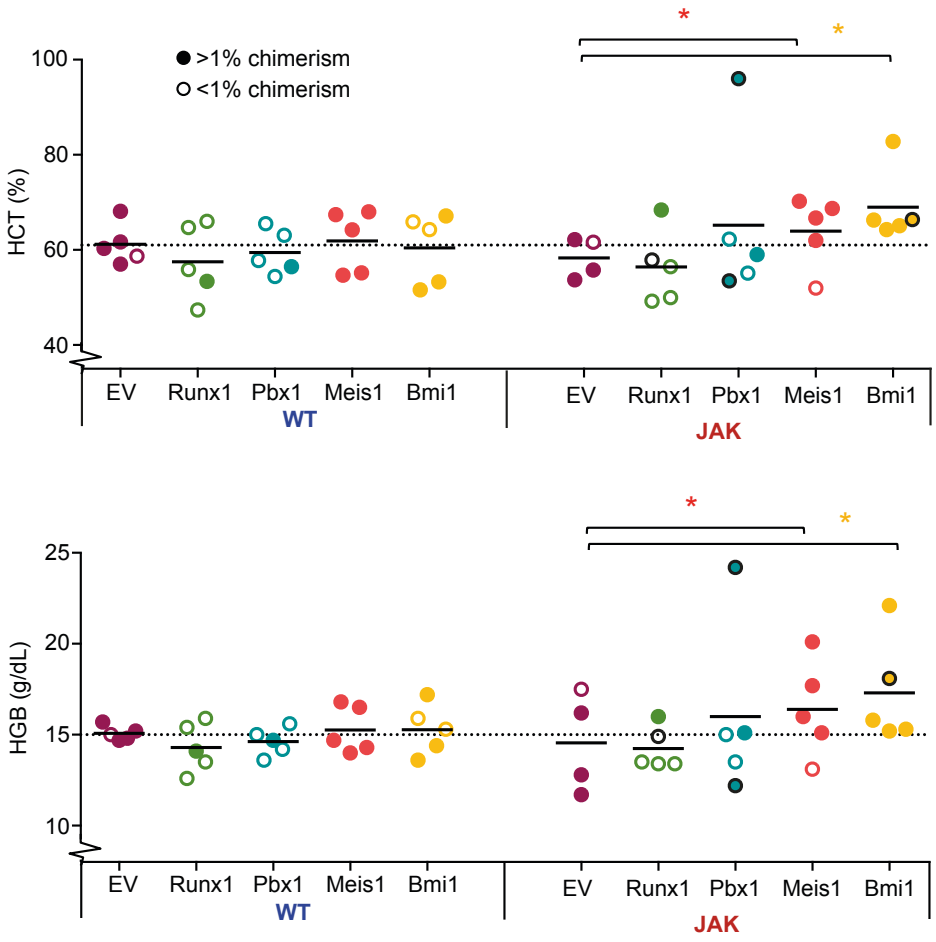


Figure 55 Haematocrit and haemoglobin in recipients of WT or JAK HOM cells overexpressing candidate genes at 20 weeks post transplantation
Filled circle represents chimerism over 1% i.e. successfully repopulated. Dark circle represents a mouse that died before 20-week time point (data from last available timepoint). Comparison of repopulated mice revealed significant elevation of HCT and HGB in recipients of JAK HOM cells overexpressing Meis1 (HGB p=0.049, HCT p=0.019), and Bmi1 (HGB p=0.050, HCT p=0.017).

5.2.3 Transplantations of JAK HOM cells overexpressing key genes can result in splenomegaly

Extramedullary haematopoiesis is a common feature of MPNs; the disease clone disrupts the BM microenvironment, disrupting trafficking and normal haematopoiesis, and driving haematopoiesis into secondary sites, like the spleen. This leads to an accumulation of cells in the spleen, resulting in enlargement of the spleen –

splenomegaly.[265] 30-40% of PV patients develop splenomegaly during the course of their disease progression.[266]

Some transplantation recipients developed splenomegaly, as determined by increased spleen length and weight (Figure 56 top). Of the mice with severely enlarged spleens (over 2cm), all were recipients of JAK HOM cells; one *Pbx1*, one *Bmi1*, and one EV. Only JAK HOM *Meis1* recipients had statistically larger spleens than WT EV recipients (length, p=0.019, weight p=0.017) (Figure 56 bottom).

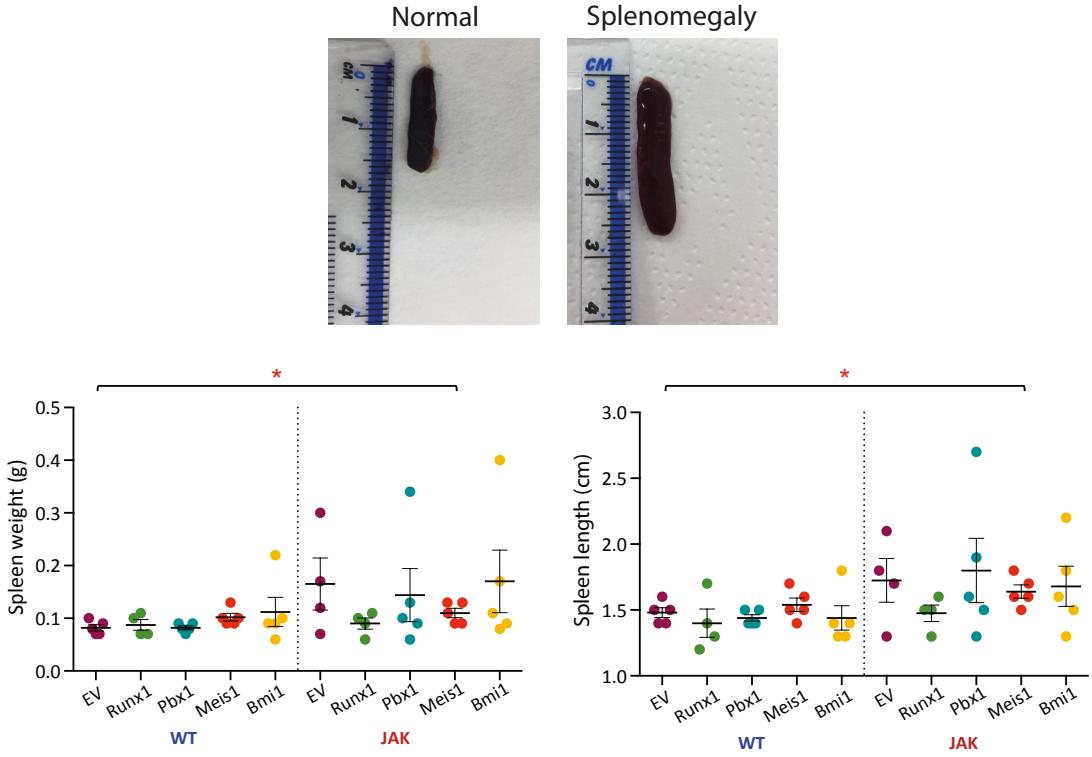


Figure 56 Assessment of spleen size in recipients of WT and JAK HOM cells overexpressing candidate genes
Representative images of a normal and an enlarged spleen. Spleen weight and spleen length only significantly larger than WT EV recipients in recipients of JAK HOM cells overexpressing Meis1 (length, p=0.019, and weight p=0.017).

5.3 JAK V617F gene expression in other settings

One of the large unanswered questions in the MPN field is if JAK2 V617F confers a self-renewal disadvantage then why/how do patients with mutant JAK2 alone present with an MPN phenotype? Relatedly, in mice receiving high cell dose transplantation, how can JAK2 alone lead to an MPN phenotype?

Gene expression analysis of the four self-renewal genes *Bmi1*, *Meis1, Pbx1*, and *Runx1* was performed to assess whether JAK mutant alone HSCs that are able to give an MPN phenotype, have achieved elevated *Bmi1, Meis1, Pbx1*, and/or *Runx1* (thereby restoring HPSC self-renewal properties) by another mechanism.

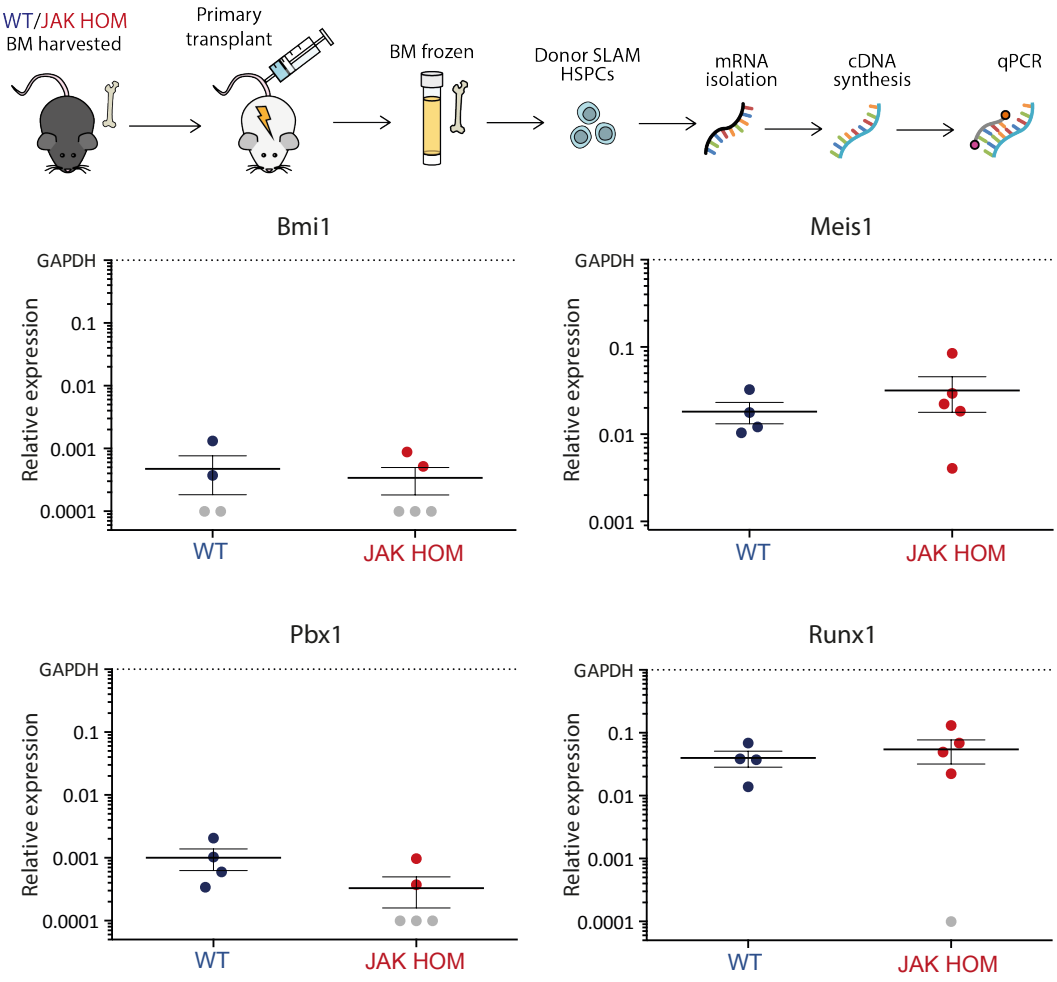


Figure 57 Gene expression of Bmi1, Meis1, Pbx1, and Runx1 in bulk HSPC post-transplantation samples

HSPCs ($CD45^+CD48^-CD150^+$) were sorted from the BM of WT and JAK HOM whole BM transplantation recipients. From these mRNA was isolated, cDNA synthesised, and gene expression assessed by qPCR using TaqMan probes. No significant differences were seen between WT and JAK HOM cells in their expression of Bmi1(p=0.705), Meis1(p=0.400), or Runx1(p=0.581) or Pbx1 (p=0.174). Relative expression calculated as $2^{-\Delta CT}$. Below threshold samples are entered at 0.0001 in grey. Error bars are SEM.

5.3.1 Bulk qPCR lacks power to detect changes in key genes in post-transplantation murine HSPCs

Donor HSPCs (SLAM) from 4 WT and 5 JAK HOM transplantation recipients were isolated, mRNA extracted, cDNA synthesised, and qPCR performed (Figure 57 top). No significant differences were seen between WT and JAK HOM cells in their expression of *Bmi1*(p=0.705), *Meis1*(p=0.400), *Runx1*(p=0.581) or *Pbx1* (p=0.174) (Figure 57). This is unsurprising for two reasons; firstly, the qPCR was performed on bulk samples, and the assayed population encompassed both stem and progenitor cells. The proportion of stem cells in the assayed population is therefore small (especially in JAK HOM animals which have a progenitor expansion), it is therefore difficult to detect changes in self-renewal genes which are predominately involved at the stem cell level. Secondly, as the changes induced by JAK2 V617F are subtle, a shift in the balance of WT HSCs, they can only be seen at single cell level; bulk gene expression analysis by microarray between JAK HOM and WT HSCs only reveals changes in cell cycle.[229]

5.3.2 Differences in expression of key genes seen in MPN patient samples

To assess whether *BMI1, MEIS1, PBX1*, or *RUNX1* expression is altered by JAK2 V617F/TET2 mutation in patients, gene expression in HSC-enriched CD34+CD38- population (~1 in 617 HSCs[267]) was assessed by qPCR from primary patient samples bearing JAK2 V617F and/or TET2 mutations. (Figure 58 top)

In agreement with the mouse model data, $CD34^{+}CD38^{-}$ cells from patients with only a JAK2 mutation had significantly lower *MEIS1* expression than non-mutant individuals (Figure 58 middle right, p=0.019), whereas in patients with both JAK and TET mutations *MEIS1* expression was not significantly different from non-mutant individuals. *BMI1* expression was significantly lower in patients with JAK2 and TET2 mutation than in non-mutant individuals (Figure 58 middle left, p=0.046) and JAK2 mutation alone patients (p=0.022), this is due to *BMI1* expression being below detection threshold in JAK2 and TET2 double mutant patients. No significant differences were observed in *PBX1* or *RUNX1* expression between non-mutant individuals and those with JAK2 or JAK2 and TET2 mutations (Figure 58 bottom). However, in agreement with mouse data, *PBX1* expression was higher in individuals with JAK2 and TET2 mutations than in individuals with a JAK2 mutation alone (p=0.026).

There are a number of reasons why differences might not be seen in the expression of self-renewal genes in patient samples. Firstly, these are bulk samples, as in the mouse it is expected that the changes evoked by JAK2 mutation would be fairly subtle, affecting changes in the balances of types of stem cell within the stem cell pool, this

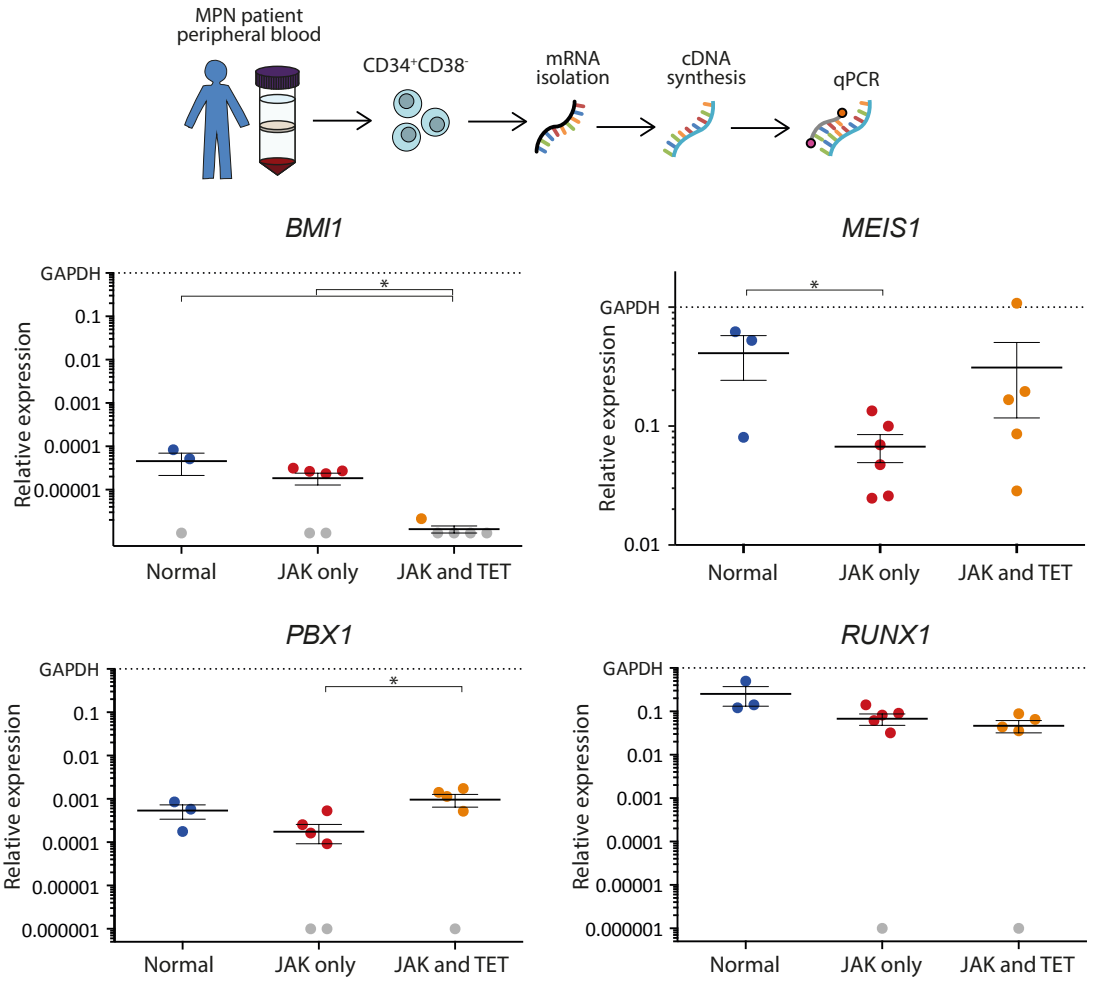


Figure 58 Gene expression of BMI1, MEIS1, PBX1, and RUNX1 in bulk CD34$^+$CD38$^-$ human samples

Schematic of experimental procedure; bulk CD34$^+$CD38$^-$ cells were sorted from normal individual or MPN patients with a JAK2 V617F mutation alone, or a JAK2 V617F in combination with a TET2 mutation. From these samples' mRNA was isolated, cDNA synthesised, and gene expression assessed by qPCR using TaqMan probes. MEIS1 expression was lower in JAK alone patients than normal individuals (p=0.019). BMI1 expression was lower in JAK2/TET2 mutant patients than JAK2 alone (p=0.022) and normal individuals (p=0.046) PBX1 expression was higher in double mutant patients than in JAK alone patients (p=0.026). No significant changes were seen in Runx1 expression. Below threshold samples are entered at 0.000001 in grey. Error bars are SEM.

kind of subtle change cannot be detected at the bulk level. Secondly, the $CD34^{+}CD38^{-}$ population is far from a pure population of stem cells (approximately 1 in 600), making it very difficult to gain a meaningful insight into expression of self-renewal associated genes which are only expressed at the very top of the haematopoietic hierarchy. Thirdly, the different mutations will also cause shifts in the ratios of different cell types within the $CD34^{+}CD38^{-}$ population, as in the bulk post-transplantation gene expression above the disruption in the balance of cell types can cause changes in the expression levels of genes in a bulk population. Fourthly, unlike mouse models, patients are not 100% mutant. A patient with mutations in both JAK2 and TET2 genes will have cells with no mutations, cells with a JAK2 or TET2 mutation alone, and cells with both mutations. A bulk assay is an average of all cell types presented, the relative proportions of each population and their contribution to the gene expression is unknown.

Due to these reasons this experiment lacked the necessary power to assess the effects of JAK and TET mutations on the self-renewal regulators of HSCs in settings like patients and post-transplantation. These findings highlight power of single cell approaches in unpicking the molecular networks of haematological malignancies (unable to see changes in bulk but are able in single cell), and sets the stage for the advent of new techniques combining single cell transcriptomics with targeted sequencing to overcome the problems connected with assessing gene expression profiles associated with different mutation in patient samples.

6

Discussion and Future Directions

6.1 Conclusions

In order for a cancer to develop and maintain itself from a single HSC, the HSC and its progeny must be able to thrive relative to the endogenous population of non-malignant HSCs. Understanding the relative self-renewal and proliferative capacity of HSCs in malignant and non-malignant populations is therefore crucial if malignant HSCs are to be effectively targeted.

As many leukaemias are complex and have many mutations, simplified disease models like MPNs are a useful starting point from which to study the effect of combinations of mutations. Where combinations of mutations are common, it becomes increasingly important to study both the mutations individually and in combination in order to gain a full appreciation of how the mutations interact to influence self-renewal and proliferation, and how the cancer develops though the sequential acquisition of a series of mutations. Combining the latest advances in single cell techniques with recently developed mouse models, it is now possible to understand more deeply the precise mechanisms involved in malignant HSC self-renewal.

In this thesis I establish the *in vitro* proliferative capacity, and *in vivo* self-renewal potential of HSCs from JAK2 V617F, TET2 knock-out and double mutant animals. I then identify the distinct molecular profiles that correlate with the observed HSC functional properties in mouse models where JAK2 V617F drives hyperproliferation,

TET2 loss increases self-renewal and double mutant HSCs have increased self-renewal and hyperproliferation. This is followed by genetic rescue experiments to identify which self-renewal regulators are critically involved in partnering with JAK2 V617F to drive a myeloid malignancy. Together, these single-cell approaches refine the molecules involved in clonal expansion of MPNs and have broad implications for deconstructing the molecular network of normal and malignant stem cells.

In chapter 3 I utilised single cell *in vitro* culture of HSCs to investigate the effect of common myeloid malignancy mutations on the cell cycle kinetics of HSCs. Comparison of WT HSC and HSCs with loss of TET2 revealed that this mutation does not affect the proliferation kinetics of HSCs. On the other hand, JAK2 V617F mutation alone caused a significant acceleration of cell cycle kinetics. When these two commonly co-mutated mutations were combined (homozygous JAK V617F with either heterozygous or homozygous loss of TET2) it was revealed that double mutant HSCs have cell cycle kinetics accelerated to the same extent as JAK2 V617F alone. This finding establishes for the first time that loss of TET2 does not restrict (or amplify) the hyperproliferative properties of JAK2 V617F in culture.

A selection of mouse models of myeloid malignancy mutations (p53, NPM1, CBP, CALR) were then assessed to determine the commonality of these effects across types of mutations. While many of these mutations have been reported to affect the cell cycle properties of cells, the study of HSCs at the single cell level still remains relatively rare and this represents the first direct comparison of this number of myeloid malignancy mutation models HSCs *in vitro*. From these data we are able to conclude that while CALR mutation affects the same pathway as JAK2 V617F and both mutations accelerate cell division kinetics, they do not have identical effects on HSCs in culture (JAK2 HSCs have quicker entry into cell cycle than CALR mutant, and CALR producing larger clones after 10 days in culture than JAK2 mutant). CBP and NPM1 mutations imposed delayed cell division kinetics upon HSCs. P53, the most common oncogenic mutation, did not cause any deviation from WT cell division kinetics. The assessment of these mutations established that the speeds exhibited by JAK and TET mutant HSCs did not represent the extents of the assay (i.e. both fast and slower cell division kinetics are possible) and established that myeloid malignancy mutation cannot be ‘grouped’ by their *in vitro* characteristics.

The effects of JAK2 V617F and loss of TET2 function were then validated in the human setting using data from MPN patient samples. The human data supported the mouse findings that JAK2 V617F drives an acceleration in cell cycle kinetics, an effect that is not hampered by the loss of TET2. These conclusions in patient data add conviction to the strength of the mouse models in recapitulating this disease.

In chapter 4, I explored the *in vivo* myeloproliferative and self-renewal capacity of JAK2 and TET2 mutant murine bone marrow. As the double mutant model had not been tested in the transplantation setting previously, it was first established that this model is capable of imparting an MPN-like phenotype upon transplantation. This was established by non-competitive transplantation of whole BM from JAK2 TET2 double mutant animals. Recipient mice displayed elevated haematocrit and haemoglobin, representative of a PV-like MPN phenotype.

Through competitive limiting dilution, it was determined that transplantation of 50,000 JAK HOM TET HOM cells or more are capable of initiating this MPN-like phenotype. This transplantation also established that (in the primary setting) that double mutant BM has equivalent repopulation and short-term self-renewal capacity to WT BM.

The gold standard for determining the self-renewal potential of a cell population is competitive serial transplantation. Competitive BM transplantation from JAK2 V617F donors confirmed previous reports that JAK2 V617F alone has a self-renewal defect resulting in the exhaustion of the donor population. Loss of TET2 had chimerism, on average, similar to WT donors. Combining JAK2 V617F with heterozygous loss of TET2 partially rescued the self-renewal defect imparted by JAK2 V617F, slowing the exhaustion of the donor population. Homozygous loss of TET2 combined with JAK2 V617F lead to rescue of the JAK2 V617F self-renewal defect, allowing serially transplantable sustained chimerism and the accompanying elevation of haematocrit and haemoglobin.

These findings, combined with the *in vitro* findings of chapter 3, allows proliferative and self-renewal characteristics to be assigned to HSCs of each mutational status; JAK2 HOM HSCs have a proliferative advantage leading to erythrocytic expansion, however this comes at the expense of reduced self-renewal, JAK HOM HSCs are less

competitive than their WT counterparts and exhaust upon competitive transplantation. Loss of TET2 does not affect the speed of HSCs proliferation nor does it have a substantial effect on self-renewal when compared with WT cells. JAK HOM TET HOM HSCs have the proliferative advantage of JAK2 V617F alone, but with self-renewal restored to WT levels by the loss of TET2.

In chapter 5, I coupled these proliferative and self-renewal characteristics with expression of genes known to be involved in HSC self-renewal, in order to define the molecular mechanisms behind the JAK2 V617F self-renewal defect. For this single HSCs from WT, JAK HOM, TET HOM and JAK HOM TET HOM mice were analysed for their expression of 45 self-renewal regulators by multiplexed qPCR. The single-cell gene expression profiling provided molecular evidence of distinct HSC states, and these profiles parallel the unique functional features of HSCs with different genotypes. As each distinct molecular profile could be identified in WT HSCs, although sometimes at a low frequency, these data suggest that WT HSCs are a heterogeneous mix of cells in distinct proliferative and self-renewing states. JAK2 and TET2 mutations disrupt the balance of HSC molecular subtypes by restricting the number of HSC states and, subsequently, via their downstream progeny, lead to an MPN phenotype. Comparison of gene expression between JAK HOM and JAK HOM TET HOM HSCs, which both share hyperproliferative properties but differ in self-renewal potential, allowed identification of key self-renewal genes under in JAK HOM HSCs; namely *Bmi1, Pbx1, Runx1*, and *Meis1*. The involvement of these genes in the self-renewal of MPN HSCs was then validated by genetic rescue experiments by lentiviral overexpression of each gene in JAK HOM HSPCs. Transplantation of these cells was met with variable engraftment success. When able to engraft, overexpression of *Meis1*, *Pbx1*, or *Bmi1* in JAK2 V617F HSPCs were able to, in some cases, elicit a myeloproliferative phenotype, and in 2 cases resulted in premature death of recipients and splenomegaly.

The expression of these four key self-renewal genes were assessed in post-transplantation JAK HOM HSPCs and MPN-patient HSPCs, however as stem and progenitor cells were assessed at the bulk level no changes in expression of these genes was able to be seen; highlighting the power of single cell approaches in uncovering the molecular networks governing self-renewal in malignant HSCs.

6.2 Discission and remaining questions

6.2.1 Driver mutations and proliferation

In this thesis the proliferation of HSCs from a number of myeloid malignancy mutation models was explored at the single cell level. Cell cycle kinetics from some of these models have previously been studied at the bulk level, however this is the first assessment of these models at the single cell level. The benefits of performing this assay at the single cell level include the guaranteed clonality of growth, and ability to see the heterogeneity in the population. As this is an *in vitro* system, this assay also has the benefit of being relatively fast; within 11 days you can compare key HSC properties such as cell cycle kinetics, survival and differentiation. How the single cell function data from each model fits into the previously published reports is discussed below.

JAK2

The data in this thesis shows that homozygous JAK2 V617F confers a proliferative advantage onto ESLAM HSCs in culture, resulting in shorter time between cell divisions and larger colony size after 10 days in culture. Previous *in vitro* reports on homozygous JAK2 V617F have established that bulk populations of lineage negative FL cells have a proliferative advantage compared to WT cells.[228] However, as this assay was performed on a bulk population it cannot be determined if the stem cells or the progenitors are hyperproliferating, and, if the HSCs are hyperproliferative, if all the HSCs are hyperproliferative or if only a subset have enhanced cell cycle properties.

In the heterozygous setting this question had been answered by Kent et al. (2013)[230] in their study of single cell division kinetics of ESLAM HSCs from heterozygous JAK2 V617F mice. This study showed that these cells have accelerated cell division kinetics when compared with their WT counterparts, and give rise to larger, more differentiated, clones after 10 days in culture.[230]

In contrast with Li et al. (2010)[227] where JAK HET had only a modest effect upon division kinetics, Kent et al. (2013)[230] show a striking effect evoked by this mutation. This disparity is likely due to the difference in the Cre drivers of these mouse models; the homozygous JAK2 V617F model (used in Li et al. 2010) is activated from foetal development (Stella-Cre), whereas the heterozygous JAK2 V617F model (used in Kent

et al.) is inducible by pIpC treatment (Mx1-Cre). pIpC treatment is an inflammatory stimulus and consequently has broad effects on the haematopoietic system, including potentially enhancing proliferation of HSPCs *in vivo* (also shown to be very leaky in same paper).[268] Due to this, induction by Stella-Cre may represent a more physiologically relevant model, and also allows for generation of homozygous mutation, which is more prominent in PV patients, although the deficiency is that it is expressed in all cells of the body rather than specifically in the haematopoietic system.

The data in this thesis uses the Stella-Cre homozygous JAK2 V617F model to confirm the conclusions of Kent at al.; at the single cell level JAK2 V617F accelerates the cell division kinetics of HSCs.

CALR

Unlike JAK HOM HSCs which have accelerated cell cycling from the first cell division, CALR HOM HSCs do not have faster entry into cell cycle, but do have a shorter time to second division, and after 10 days in culture produce a larger proportion of large colonies than even JAK HOM HSCs. This coupled with the reported normal self-renewal of CALR HSCs (neither increased nor decreased self-renewal potential as determined by serial transplant)[201] implies that the speed of exiting the stem cell state (i.e. time to first division) is not changed in CALR HOM HSCs but once cycling they have substantially elevated proliferative potential.

While JAK2 V617F and CALR del52 mutations both affect the same cellular pathway (i.e. the JAK/STAT signalling pathway) and have broadly similar effects in that they both cause hyperproliferation, they do not perform identically in culture (or *in vivo*). This could be due to differences in mouse models – the JAK2 model uses the entire human JAK2 gene whereas the CALR model combines the human mutant c-terminus with the murine gene, JAK is Stella-Cre (activation from embryo) while CALR is Mx1-Cre (induced by pIpC treatment, which as mention previously has been proposed to accelerate HSPC cycling kinetics)[268] – or it may be due to the differences in their mode of driving hyper-activation of the JAK-STAT signalling pathway.

CBP

The CBP mutant mouse used in this thesis has been reported to have a myeloid (GM) differentiations bias and self-renewal defect, as determined by serial competitive transplantation.[239] This model has been reported to have a decrease in LSK number (progressive over time) and increase in downstream myeloid progenitors (Lin^- IL-7^- c-Kit^{hi} $Sca1^-$). CBP null HSCs share a number of characteristics with JAK2 V617F HSCs; a self-renewal defect, expansion of downstream progenitors, and a myeloid cell bias. Despite these similarities, CBP mutant HSCs behave in the opposite manner to JAK2 mutant HSCs in culture; CBP mutant HSCs have delayed cell division kinetics and smaller average clone size after 10 days in culture.

There are several possible explanations for this difference. While in JAK2 mutant HSCs the self-renewal defect is suggested to be due to increased proliferation and differentiation of these cells (prioritised over slow-dividing quiescent HSCs), this is by no means the only way of achieving a reduced stem cell number and activity. HSCs could also be lost due to death or direct differentiation, or reduced division frequency meaning the stem cell population size is not maintained. In CBP null animals the progressive loss of stem and progenitor cells with impeded division kinetics could be due to any of these reasons. The myeloid skewing reported in CBP null animals[239] might not represent a myeloid bias of progenitor cells but instead a reduction in population sizes of other cell types as stem and progenitor cells exhaust.

TET2

TET2 has established roles in self-renewal, however few studies (one) look at the proliferation of TET2 stem and progenitor cell in culture, instead many focus on transplantation and CFCs which can be used to study self-renewal. TET2 knock-out LSK cell have been reported to hyperproliferate *in vitro*[211]; when 500 LSK cells were cultured for 7 days (liquid culture with 4 growth factors - mSCF, mIL-3, IL-6 and EPO) they generated more total cells than WT, and when replated in CFCs, TET2 knock-out cells were able to generate more colonies. As this is a bulk culture it cannot be known if HSCs or progenitors are responsible for this enhanced proliferation.

The data in this thesis shows that TET2 knock-out alone does not impact division kinetics of ESLAM HSCs in culture compared to WT HSCs, when cultured with the

cytokines SCF and IL-11. The different cell types (ESLAM vs LSK) could be responsible for this difference (i.e. in the LSK culture, HSCs might be proliferating at normal speed but progenitors hyperproliferating), or the difference in the supplemented cytokines might be responsible. A potential future experiment would be to compare the division kinetics of single cell ESLAM HSCs and MPP1 progenitors ($Lin^-Sca1^+c\text{-}Kit^+CD48^-CD150^+CD34^+FLK2^-$) in medium supplemented with SCF, IL-3, IL-6 and EPO.

NPM1

NPM1 mutant HSCs have been reported to have increased self-renewal, like TET2 knock-out HSCs, and a myeloid bias. Despite this myeloid bias, when cultured in conditions supporting myeloid differentiation NPM1 mutant HSCs had slower cell division kinetics and consequently gave rise to on average smaller colonies that WT HSCs. However, these mice also have a reported enhancement of HSC self-renewal,[240] and since accelerated cycling is associated with exiting the stem cell state and differentiating,[269] it may be that NPM1 HSCs retain their self-renewal state by dividing more slowly and that downstream progenitors give rise to the myeloid expansion.

P53

In this thesis no difference in cell division kinetics was seen *in vitro* when compared with WT HSCs. As P53 is known to affect self-renewal, rather than proliferation, through its interaction with EZH2 and subsequent effect on epigenetic regulation (and self-renewal potential is not reported in this assay), it is therefore not surprising that P53 knock-out HSCs act similarly to TET2 knock-out HSCs (which also have increased self-renewal with no proliferative change). However, loss of P53 has been reported to accelerate cycling of HSCs *in vivo;* BrdU (bromodeoxyuridine) staining has shown that *in vivo* an increased proportion of $p53^{-/-}$ HSCs are cycling relative to WT cells, two days after administration of BrdU 60% of $CD34^-$LSK cells from $p53^{-/-}$ animals contained the stain, vs 30% in WT animals.[261] This disparity between *in vitro* and *in vivo* data could be the consequence of a lack of microenvironmental factors *in vitro*, which may be required for hyperproliferation of P53 knock-out cells. It could also be that it is not the HSCs with accelerated kinetics but a downstream progenitor; the cells assessed for BrdU staining *in vivo* were $CD34^-$LSK cells which are reportedly 21%

HSCs[10], however we might still expect to see this hyperproliferation in culture in the day 10 colony size read out.

This model confirms that not all oncogenic mutations affect HSC characteristics *in vitro* and highlights the importance of combining *in vitro* assays with *in vivo* experiments to validate the observed effects are applied in the complex network/environment of the haematopoietic system. Nonetheless, this set of experiments exemplify that valuable insight can be gained from studying HSCs *in vitro,* for example the finding that CALR and JAK2 mutations do not cause the same affect on HSC proliferation despite effecting the same pathway. These experiments have the additional benefits of being much higher throughput than *in vivo* assays and can therefore be a quicker and easier way to identify biological features associated with mutations.

6.2.2 Differences between effects of JAK2 and TET2 mutation in mouse and human

When the results in mouse models were compared with those from human patient samples, some findings were borne out, others were not. In terms of proliferation, JAK V617F hHSPC behaved similarly to murine HSCs, with faster entry into cell cycle and reduced time to subsequent divisions. However, this hyperproliferative effect did not extend to day 10 colony size, with no significant increase in proportion of large colonies from JAK2 V617F hHSPCs. The difference in effect of JAK2 V617F on day 10 colony size between human and mouse samples could be due to the more mixed starting population in the human samples (i.e. hHSPC rather than mouse HSC) having lower/finite proliferative capacity. Moreover, the lack of effect of hyperproliferation on day 10 clone size could be a consequence of the culture conditions for human HSPCs being less well optimised than those for mouse. The reduced efficiency of human culture is evident in a number of ways; firstly compared with mouse, a higher proportion of sorted hHSPCs do not give rise to a colony which could be due to the mixed starting population, more committed progenitors lacking the capacity to form a colony under these restricted cytokines/environmental stimuli. Secondly, the cells which are able to give rise to colonies grow more slowly than in mouse colonies[269] and a much smaller proportion are able to reach medium or large size, this is compounded by the fact that some human colonies can begin dying after 9-14 days in liquid culture. While the culture conditions are not as well optimised for hHSPCs the early division

kinetics (i.e., the time to first and second division) are recapitulated in the human JAK2 V617F cells in line with the mouse data, demonstrating that cells with a JAK2 activating mutation have faster cell cycle kinetics than those without.

Human HSPCs with a TET2 mutation proliferate in a similar manner to WT cells, but unlike TET2 knock-out mouse HSCs, after 10 days in culture a larger proportion of the clones from TET2 mutant hHSPCs contain fewer than 50 cells compared to non-mutant clones. While the human assay does not commonly produce medium and large colonies (due to the slower initial cycling of hHSPCs), the change in the ratio of very small to small colonies seems to represent a real change in proliferative capacity of TET2 mutant HSPCs. This could be a function of the change in the ratio of stem to progenitor cells in the HSPC population; TET2 give a self-renewal advantage giving an increase in the HSC population, HSCs being quiescent are likely to divide more slowly/exit quiescence slower and consequently give rise to smaller colonies. Alternatively, the difference could be the result of the fact that in the MPN patients TET2 mutations are commonly on a single allele, whereas in the mouse model TET2 is completely knocked out. Comparison of TET2 mutant mice (loss of catalytic activity, not whole protein) with complete TET2 knock-out mice has revealed that TET2 has non-catalytic roles in haematopoiesis.[270] TET2 mutation (loss of catalytic activity) leads to an expansion of HSPCs and enhanced colony formation compared to WT, but neither the HSPC expansion nor colony forming capacity is enhanced to the same degree as in TET2 complete knock-out. The mice also differ in the diseases they develop with TET2 mutant animals developing myeloid malignancies while complete loss develop both myeloid and lymphoid malignancies.

It is also formally possible that as humans and mice diverged evolutionarily a very long time ago, not all protein functions have been entirely conserved, TET2 could have roles in human that are absent in mouse. However, this would be unlikely to affect its role in leukaemogenesis since in both cases loss of function mutations are associated with myeloid malignancies.

As mentioned above, neither the starting cell population nor the culture conditions for human cells used here are as well optimised as they are for murine HSCs. This is an area of ongoing improvement, with recent advances including the combining of

coculture methods (on MS5 stromal cells) with addition of various growth factors and cytokines.[28] However even if human HSC culture were fully optimised, HSC isolation from PB is far from ideal. Using the current best sorting strategy from cord blood, only ~1 in 10 sorted cells is an HSC,[29] and this is much lower in peripheral blood or bone marrow obtained from patient samples. Until there are advances in the isolation of human HSCs, the study of HSCs will remain largely clouded by non-HSC contaminants, making interpretation of single cell data extremely challenging. Using current methods, studying HSPC fractions permits large effects (that effect both stem and progenitor cells) to be seen, like that of JAK2 on proliferation, however more subtle changes, such as affecting HSC self-renewal at a modest level, are not resolvable. Therefore, until this technical barrier is surmounted, the study of HSC mechanisms will continue to rely on mouse models.

6.2.3 Order of mutation acquisition

Like in mouse, human JAK TET double mutant cells act similarly *in vitro* to cells carrying a JAK2 V617F mutant alone, with accelerated cell division kinetics and increased proliferation. This allows the same conclusion to be drawn in the human setting as in mouse; namely that JAK2 V617F hyperproliferation is not restricted by a concomitant loss of function TET2 mutation. It should be noted that in the patients the mutations are acquired sequentially, either JAK2 first or TET2 first, whereas in the mouse models used in this thesis, both mutations are germline. This could have some consequences, since the order of mutation acquisition has been shown to influence both the clinical features of disease (e.g. disease subtype and risk of thrombosis), as well as aspects of HSPC biology (e.g. CFC generation capacity and balance of progenitors).[190]

The order of mutation acquisition represents an important and understudied aspect of combining mutations. A proposed future experiment is to model the order of JAK2 and TET2 mutation acquisition in mice. Modelling the order of mutation acquisition in mouse rather than human has a number of advantages; firstly, it is possible to isolate HSCs at a much higher frequency in mouse than in human. Secondly, reporters can be introduced alongside mutations, allowing for easy prospective isolation of cells with different genetic alterations (in humans the mutations present can only be retrospectively identified). This also widens the scope of assays which can be performed on HSCs of different mutational status as the assays no longer need to result

in the growth of a colony for retrospective genotyping. Thirdly, patients are influenced by a huge number of external factors (differences in lifestyle, smoking, diet, etc) which cannot be controlled for, in mouse it is possible to control for these factors as well as others, such as time of acquisition of first mutation and length of time to acquisition of second mutation.

Even more practically, the order of mutation acquisition could be modelled by lentiviral overexpression of JAK2 V617F in TET2 knock-out HSCs, and by shRNA (short hairpin RNA) knock-down, or CRISPR knock-out of TET2 in JAK2 V617F HSCs. Alternatively, the order of mutation acquisition could be modelled by crossing the current models with inducible models. The crossing of TET2 knock-out mice with an inducible JAK2 V617F model (e.g. pIpC inducible [271]) would allow the modelling of 'TET first' order, and crossing the JAK V617F mouse with an inducible TET2 knock-down model,[272] would act as the 'JAK first' model. The resultant HSPCs could be assessed in many of the same ways as the mouse models in this thesis; *in vitro* for their time between cell divisions and subsequent day 10 clone size and composition, and in transplantation for their myeloproliferative phenotype (order affects PV vs ET) and self-renewal capacity, and for differences in gene expression between the orders (and similarity to single mutant HSCs).

6.2.4 Transplantation variability in recipients of TET2 bone marrow

In competitive transplantation, it was seen that recipients of TET2 knock-out BM had highly variable chimerism (section 4.3). From published literature TET2 knock-out HSCs have a self-renewal advantage,[135,210,211] resulting in high chimerism when competitively transplanted. The low chimerism in some of the recipients in this thesis is therefore somewhat perplexing. One potential technical cause of this could be sub-optimal transplantation injection resulting in a lower dose of cells being transplanted but there are a few reasons why this is unlikely to be the cause. Firstly, the number of donor and competitor cells is within the normal range, implying that transplantation itself was performed adequately. Secondly, the same variability in chimerism is seen in TET HOM alone and JAK HOM TET HOM recipients but not in WT recipients, consequently it seems to be a phenotype associated with loss of TET2 and not one of the experiment in general.

A further potential explanation for the difference between this thesis and previous studies might lie in the different transplantation recipient strains used. Previous transplantation of TET2 knock-out cells have used lethally irradiated Bl6 mice as recipients,[135,210,211] whereas in this thesis the recipients were sub-lethally irradiated W41 mice. W41 mice have a partial loss of function in the SCF receptor c-Kit, this mutation results in an HSC defect.[95,273] This defect allows these mice to act as transplantation recipients after receiving a sub-lethal dose of irradiation as the recipients own HSCs are less competitive and therefore outcompeted by transplanted donor cells. As W41 mice only receive sublethal irradiation they have reduced morbidity and mortality compared with other models that require lethal irradiation and are consequently ethically preferable. A potential reason why transplantation of TET2 knock-out cells into W41 recipients might differ from Bl6 recipients could be due to differences in extramedullary haematopoiesis. TET2 knock-out animals have been reported to have an increase in extramedullary haematopoiesis, and the increase in self-renewal seen in TET2 knock-out mice could be the result of, not of an increased intrinsic self-renewal capacity, but an extension of the niches in which an HSC is able to self-renew. While sublethal irradiation of W41 mice is sufficient to allow engraftment of transplanted donor cells in the BM, it may be that sub-lethal irradiation is insufficient for evacuation of extramedullary haematopoietic sites, meaning TET2 knock-out HSCs would be unable to occupy all the same sites in W41 recipients as in Bl6 resulting in reduced self-renewal and consequently lower chimerism. Supporting the hypothesis that splenic sites may be occupied in sub-lethally irradiated W41 mice, it has been reported that other c-Kit mutations ($Kit^{W\text{-}sh}$) cause extramedullary haematopoiesis in the spleen.[274] Alternatively, it may be that the difference between lethal and sub-lethal results in other microenvironmental and inflammatory changes that effect TET2 knock-out HSCs. A potential future experiment would be to do an assessment of the impact of different recipient strain and irradiation doses on TET2 self-renewal capacity.

6.2.5 Functional assessment of combinatorial mutations

While combinations of mutations are common in patients, and the number of double mutant mouse models is increasing, there are limited reports on the functional effect on HSCs of co-mutation of JAK2 and TET2. Primary competitive transplantation has been performed on the only other published JAK2 V617F/TET2 knock-out mouse model.[225]

This group transplanted LSK cells from each genotype alongside WT LSK and reported, in agreement with the data in this thesis, that loss of TET2 in combination with JAK2 V617F improved chimerism when compared with JAK2 V617F alone. This study did not include secondary transplantation so robust conclusions could not be drawn about the long-term self-renewal of double mutant HSCs, but the data in this thesis would support it being robust in serial transplantations.

The findings of this thesis and Chen et al., that loss of TET2 function improves JAK2 V617F HSC function in transplantation, is also supported by xenotransplantation findings. When $CD34^+$ cells from JAK2 V671F MPN patients low allele burden (less than 50% in granulocytes) were transplanted into nonobese diabetic/severe combined immunodeficiency (NOD/SCID) mice, cells were unable to engraft in recipients.[275] However when xenotransplantation was performed from patients with both JAK2 and TET2 mutations, donor $CD34^+$ cells showed enhanced repopulating capacity compared with JAK2 mutation alone.[276] This would suggest that patient JAK2 V671F HSCs do not display a robust proliferative advantage but instead have reduced engraftment and limited self-renewal capacity,[277] in agreement with some mouse models.[228] This also supports mouse model findings that loss of TET2 improves the ability of JAK2 V617F HSC to sustain long-term engraftment in transplantation recipients. The inability of JAK2 V617F patient samples to reliably engraft in mice is also one of the key reasons why mouse models have been so heavily used to study this mutation.

While this thesis focused on the effect of combinations of mutations on HSCs, this is by no means the only cell population likely to be affected by the mutations. It has previously been reported that both JAK V617F and TET knock-out mutant mice have expansion in their progenitor compartments.[135,210–212,228] Meanwhile JAK V617F/TET2 knock-out have not been reported to have an expansion in progenitors in BM (but do have expansion in spleen[225]). The data in this thesis agrees with these reports. Unpublished 10X transcriptomic data has further suggested that both JAK2 V617F and TET2 knock-out mice have increased cell cycle kinetics in multipotent progenitors (S Watcham, personal communication) which could explain these expanded population sizes. A potential future experiment would be to determine if immature progenitors (MPP1, $Lin^-Sca1^+c\text{-}Kit^+CD48^-CD150^+CD34^+FLK2^-$) from JAK HOM alone, TET HOM alone and double mutant animals have different cell cycle kinetic *in vitro*. This

would be assessed by single cell sorting of MPP1s into erythroid permissive media (and the standard media used in *in vitro* experiments in this thesis), and tracking the cell division kinetics, and assessing the clone size and composition after 14 days by flow cytometry.

6.2.6 Unpicking the molecular states of HSCs – molecular drivers of cellular properties

In the assessment of HSC gene expression, approximately 25% of the JAK HOM HSCs had reduced expression of the key self-renewal regulators *Bmi1, Meis1, Pbx1,* and *Runx1*. One potential explanation for the increased proportion of cells lacking HSC self-renewal regulators in the JAK2 homozygous mouse is that the phenotypic HSC gate (ESLAM) captures a different proportion of HSCs compared with contaminating progenitors than its wildtype littermates (e.g. there are more contaminating non-HSCs). Several lines of evidence suggest this would not be the case, including the similar reduction in HSC number using an alternative sorting gates ($Lin^-Sca1^+c\text{-}Kit^+CD34^-Flk2^-$ [164]) and the observation that HSCs with similar molecular programs are present in WT mice (albeit at a lower frequency), but not in the TET knock-out or double-mutant cells. In either case, as this molecular study represents a static picture taken at a single developmental stage (3-4 months) and does not give any information regarding the dynamics of the HSC compartment, it is not currently possible to resolve the relationship of impaired to unimpaired HSCs, and we cannot yet ask questions of primacy or relatedness.

The molecular signature of TET HOM HSCs partially overlaps with JAK HOM TET HOM HSCs but neither overlap with JAK HOM HSCs. The genes with reduced expression in TET HOM HSCs (that are not reduced in JAK HOM TET HOM) are *Vwf, Gata3, Prdm16,* and *Tal1*. Previous gene expression profiling of TET2 knock-out cells have reported significantly downregulation of *Hoxa9* in LSK cells and *Gata2* in Lin^- cells.[270] This study agrees that *Gata2* expression is reduced in cells with loss of TET2 (with or without JAK2 V617F), but reduction in *Hoxa9* is not observed. Interestingly, many of the genes with reduced expression in models with loss of TET2 (TET HOM and JAK HOM TET HOM) are associated with higher expression in 'true' HSCs (gene expression signature of overlapping region of different HSC gating strategies[27]), i.e. *Gata2, Gata3, Gfi1b, Itga2b, Prdm16, Procr, Tal1*, and *Vwf.* There are many processes

central to maintenance of an HSC "state", including self-renewal, quiescence, avoiding apoptosis, maintaining multilineage capacity, and avoiding differentiation. Loss of TET2 has been linked to increased self-renewal, but this may not represent an increase in total 'stem-ness'. The difference in gene expression between TET mutant HSCs and 'true' HSCs might be related to some of these other processes. Indeed, *Vwf, Itga2b* (CD41), and *Gfi1b* have been linked with megakaryopoiesis, *Tal1* has been implicated in formation, but not maintenance, of HSCs, and *Prdm16* has been implicated in avoiding apoptosis and division (deletion enhances apoptosis and cycling of HSCs).

Gene expression profiling has previously been performed on bulk LSK populations from JAK2 mutant, TET2 knock-out and double mutant mice by Chen et al. (2015)[225] by whole genome Illumina MouseRef-8 v2.0 gene expression arrays (microarray). This yielded identification of 17 differentially expressed genes between the 3 groups. The authors then employed gene set enrichment analysis to look for more subtle changes in the transcriptome evoked by these mutations and found that targets of STAT5 signalling were differentially expressed in JAK2 mutant and double mutant LSK cells (but not TET2 alone), and an HSC self-renewal signature was significantly enriched in TET2 knock-out and double mutant (but not JAK2 alone). While this is supportive of the increased self-renewal of HSCs with a TET2 mutation, the lack of detection of individual differentially regulated genes means that the individual genes responsible for the defect cannot be identified and validated without using single cell approaches.

None of the differentially regulated genes identified in this study are implicated in self-renewal and the lack of differential expression of self-renewal genes between the JAK2 and double mutant cells is unsurprising for three reasons. Firstly, the assayed cell population, LSK, is a very broad population, while this gating encompasses all HSCs it also includes many other cell types and HSCs are at a very low frequency (~10% [278]). Changes affecting gene expression in HSCs are therefore unlikely to be observed in this population. Secondly, even if a more purified population of stem cells was assayed, changes in self-renewal regulators might not be seen as in this case since a bulk gene expression assay was performed. From the single-cell gene expression data in this thesis, the majority of JAK2 V617F have 'normal' expression of self-renewal regulators and only a subset of JAK2 mutant HSCs that have reduced expression of key self-renewal genes drive the differences observed. Subtle changes in the relative balance of

'types' of HSC therefore cannot be detected at the bulk level and are further supported by the presence of WT cells across the entire molecular landscape (i.e., all molecular subtypes are possible, but a mutation selectively enriches/depletes one type over another). Thirdly, this paper uses a 'whole genome' microarray approach, where amplification occurs prior to hybridisation and could result in biases in genome coverage and representation bias,[279] in the presence of mutations like JAK2 V617F which have such a large effect on proliferation. This takes over the gene signature and it cannot be determined if the absence of self-renewal genes in the gene expression profile is due to down regulation of these genes, or if the transcriptome is being flooded with proliferating genes, obscuring the presence of lowly expressed self-renewal genes.

With new technologies improving the depth of sequencing and reducing the cost, unbiased approaches are becoming more popular while targeted approaches fall out of fashion. In contrast to this, the data in this thesis rather demonstrates the power of targeted approaches for unravelling molecular networks associated with individual properties.

6.2.7 Implications for other genetic drivers

From the clustering and PCA, key self-renewal genes under-expressed in JAK HOM HSCs were identified; *Bmi1, Pbx1, Runx1*, and *Meis1.* The other change in gene expression between JAK HOM HSCS and JAK HOM TET HOM HSCs is the reduced expression of *Gfi1b* and *Dnmt3a* in double mutant HSCs. DNMT3A, like TET2, is involved in DNA methylation (as discussed in section 1.10.2) and loss of function DNMT3A mutations are common in patients with MPN, AML and ARCH. Knock-out of DNMT3A in mice leads to expansion of the HSC compartment and increased self-renewal of HSCs.[64] The finding of reduced expression of *Dnmt3a* in TET2 knock-out HSCs implicates DNMT3A in the mechanism of TET2s impact on self-renewal. A potential future experiment is to validate this interaction. As *Dnmt3a* is itself a commonly co-mutation with JAK2, knock-down of *Dnmt3a* in JAK HOM HSC is not an appropriate means of validating the effects of TET2 mutation. DNMT3A could also be overexpressed in JAK HOM TET HOM HSCs, potentially reversing the self-renewal advantage, this is a potential future experiment. This experiment could be complicated by the fact that JAK HOM TET HOM, like TET HOM alone, has highly variable chimerism (either very high or very low) making changes in self-renewal difficult to

confirm unless they are very striking. Another potential future experiment would be to cross JAK2 V617F mice with a DNMT3A knock-out model, and to repeat the gene expression with another combination of mutations, to see if the same genes lead to the restoration of the JAK2 self-renewal defect, or if a variety of self-renewal genes are able to restore the defect.

6.2.8 A role for the *Meis-Hox* axis in MPNs

MEIS1 and PBX1 control gene expression through their interaction with HOXA9.
HOX proteins are best known for their involvement in anterior-posterior patterning in drosophila,[280] and during development play a fundamental role in the regional identity of a wide range of tissue types across diverse taxa. HOX proteins are DNA binding proteins, able to bind hundreds to thousands of target genes, the specificity of HOX protein binding is assisted by co-factors, and consequently these proteins have an important role in gene expression. *Hoxa9* is the most highly expressed HOX gene in the haematopoietic compartment. MEIS1 and PBX1 are HOXA9 cofactors; HOXA9 can form a dimer with PBX1 or a trimer with PBX1 and MEIS1.[281]

Hoxa9, Meis1 and *Pbx1* have all been implicated in myeloid malignancies, especially AML. *HOXA9* overexpression is the strongest predictor or poor prognosis in AML.[88] *PBX1* fusion (*E2A-PBX1*) is a common initiating mutation in ALL.[282] *MEIS1* has been implicated in AML, both as a transcriptional target of MLL-fusion proteins,[103] and in cytogenetically normal AML, *MEIS1* expression is associated with poor prognosis.[102]

MEIS1, PBX1 and HOXA9 have all been associated with maintenance of self-renewal in HSCs. Inactivation of PBX1 leads to progressive loss of HSCs, reduced quiescence, and defective self-renewal as determined by serial transplantation. Likewise, loss of MEIS1 leads to a reduction in HSC numbers, reduced quiescence, and reduced colony formation capacity, and loss of MEIS1 in combination with MLL oncogenes removes self-renewal advantage required for myeloid transformation.[283] Overexpression of HOXA9 results in expansion of HSCs and early progenitors, leading to myeloproliferative phenotypes in mice.[88]

From the overexpression of *Pbx1* and *Meis1* in this thesis, neither gene drives a strong self-renewal advantage. As these gene products form a complex with HOXA9, it may

be that *Hoxa9* must also be overexpressed for any effect to be visible. This is a potential future experiment; overexpressing *Meis1* or *Pbx1* in combination with *Hoxa9* in JAK HOM HSCs and assessing for restoration of self-renewal defect. Combining overexpression of *Hoxa9* and *Meis1* has previously been shown to induce AML; retroviral over-expression of *Hoxa9, Meis1* or *Pbx1* alone failed to transform mouse BM cells, but combining overexpression of *Hoxa9* and *Meis1* (but not *Hoxa9* and *Pbx1*) was able to induce AML in transplantation recipients, resulting in premature death.[284] As MPNs can transform to AML, it stands to reason that they share some molecular characteristics with this more aggressive disease, it might be that high level overexpression of *Hoxa9* and *Meis1* drives AML, but that lower level increases of the same genes plays a role in MPNs.

Another potential future experiment could be to disrupt the interaction between PBX1 and MEIS1 in double mutant cells; Pbx-regulating protein-1 (PREP1) and MEIS1 compete for binding of PBX1,[285] PREP1 has been reported to reduce *Meis1* transcriptional activity and inhibit tumorigenicity.[286] Over-expression of *Prep1* in JAK HOM TET HOM HSCs would determine if disruption of the MEIS1/PBX1 complex is sufficient to disrupt self-renewal, potentially identifying a druggable pathway for the treatment of MPN patients.

6.2.9 Resolution of the impact of *Runx1* on self-renewal

The effect of RUNX1 on self-renewal has been the subject of numerous contradictory reports. RUNX1 forms the DNA binding subunit for the RUNX1/CBFβ transcription factor complex and is essential for haematopoiesis.[287] *RUNX1* chromosome abnormalities are common in AML and ALL (e.g. *RUNX1-CBFA2T1, ETV6-RUNX1, RUNX1-MECOM (MDS1-EVI1*)). Non-translocation mutations in *RUNX1* have been reported in ALL, AML, MDS, CMML. *RUNX1* is one of the most commonly mutated genes in MDS, mutated in 10% of patients. In general, RUNX1 mutations are loss of function.[287]

There have been numerous opposing reports on the effect of loss of RUNX1 on HSC self-renewal and stem and progenitor cell number. Some stating that loss of RUNX1 increases self-renewal and stem cell number (limiting dilution[288]), some that loss of RUNX1 decreases self-renewal capacity (in primary transplantation[288]), and others still

that RUNX1 has little effect on self-renewal (as determined by serial transplantation[115]). The results of this thesis show that overexpression of *Runx1* has a negative effect upon engraftment capacity in both WT and JAK HOM cells. This accords with a previous study of *Runx1* overexpression, which reports that transplantation of BM cells overexpression of *Runx1* could not engraft mice,[288] allowing the conclusion that RUNX1 negatively regulates HSCs.

The lack of engraftment from *Runx1* overexpressing cells could be a result of differentiation or death of the transplanted cells. As loss of function RUNX1 mutations cause a differentiation block,[289] it could be hypothesised that *Runx1* overexpression could promote differentiation. RUNX1 has been reported to have roles in erythropoiesis; during early erythropoiesis *Runx1* levels decline[290] and mutant *RUNX1* reduces Burst-Forming Unit erythrocytes (BFUe) formation in CFC (from $CD34^+$ cord blood cells).[289] Expression of *Runx1* in BFUe prevents *Pu.1* down-regulation and blocks terminal erythroid differentiation.[290] Therefore, reduced *Runx1* expression in JAK HOM HSCs is suspected to be related to erythroid differentiation rather than related to self-renewal properties.

This raises another question of why reduced *Runx1* expression associated with JAK HOM alone rather than JAK HOM TET HOM which shares the pro-erythroid phenotype? As just mentioned RUNX1 is involved in the control of PU.1 (gene *Sfpi1*). *Sfpi1* was one of the genes identified as being under expressed in JAK HOM cells when compared with WT cells (Figure 46) but was not associated with any induvial genotype when the four genotypes were compared together (Figure 44). This allows the conclusion that *Sfpi1* is not responsible for the JAK HOM self-renewal defect and likely plays a role in erythroid differentiation. Double mutant cells have found a means of achieving low *Sfpi1* expression without low *Runx1* expression therefore permitting erythroid differentiation. These finding in the role of *Runx1* shows the importance of functional validation of gene expression data.

6.2.10 *Bmi1*; a therapeutic target for MPNs?

Polycomb group proteins form two classes of complex that repress gene transcription; polycomb repressive complex 1 (PRC1) and 2. PRC1 ubiquitinates the c-terminus of histone H2A and thereby represses target genes. BMI1 is a component of PRC1. A

number of the components of PRC2 have been reported as being mutated in myeloid malignancies such as EZH2, as well as interacting protein such as ASXL, however the subunits of PRC1, including BMI1, have not been reported to be mutated in haematological malignancies. Although not mutated in haematological malignancies, BMI1 has been suggested as a biomarker in haematological malignancies with higher levels of *Bmi1* transcript correlating poor prognosis and more severe disease in myeloid malignancies (MDS, AML, CML) and lymphoid malignancies (ALL, CLL, mantle cell lymphoma).[291]

Both loss-of-function and gain-of-function studies have revealed a central role for BMI1 in the self-renewal of HSCs (loss of *Bmi1* gene induces a defect in self-renewal, while over-expression promotes self-renewal).[50] The strong effect of BMI1 on self-renewal accords with the results of this thesis where *Bmi1* overexpression had the most success at restoring the JAK2 V617F stem cell defect (as determined by number of positively repopulated mice).

In breast epithelial and breast cancer cell lines histone deacetylase (HDAC) inhibitors have been shown to downregulate *BMI1* expression.[292] HDAC inhibitors have shown potential as a cancer treatment across many cancer and tissue types (now combination therapy looking more appropriate due to the extensive effects of HDAC inhibitors on many biological processes in non-cancerous cells). A potential future experiment is to perform preliminary experiments to determine the feasibility of HDAC inhibitors as a differentiation therapy for MPNs. This would involve treating JAK HOM TET HOM double mutant cells with HDAC inhibitors to determine if this is sufficient to reduce *Bmi1* expression and consequently disrupt the self-renewal potential of these cells and cause graft exhaustion.

6.2.11 Emerging areas in the study of gene expression and clonal expansion for MPN patients

Assessment of the expression of *BMI1, MEIS1, PBX1* and *RUNX1* in human HSPCs from MPN patients, did not reveal any changes in these genes. As mentioned in the results there are a number of reasons why these changes might not be visible by profiling human samples in this way (bulk samples, very few HSCs in population, mixed genotypes, etc). The goal of dissecting the molecular networks governing

malignant HSC self-renewal is to apply the findings to treatment of patients with cancer. It is therefore desirable to find a more appropriate and reliable means of validating findings, such as those in this thesis, on patient samples. In recent years two new approaches have been published which are certain to have large impacts upon the study of haematological malignancies in patient samples.

Giustacchini et al. (2017)[293] developed a new protocol coupling single cell transcriptomics with detection of the BCR-ABL transgene in CML patients by multi-plexing *BCR-ABL*-specific primers at the reverse transcription and amplification steps of the scRNAseq protocol. This technology allows sensitive and specific detection of the *BCR-ABL* mutation simultaneously with unbiased whole transcriptome analysis of the same HSC, thereby permitting a molecular comparison of mutant to non-mutant HSCs in the same individual. After validating the nested priming approach for genotyping cells from the transcriptome, this study compared HSCs from normal individuals to BCR-ABL^{+} HSCs and BCR-ABL^{-} HSCs from CML patients.

Perhaps the most interesting aspect of the Giustacchini et al. study was their comparison of normal HSCs to non-mutant HSCs in the CML patient which showed striking differences in the gene signatures; non-mutant HSCs from CML patients had higher expression of genes associated with microenvironmental factors IL-6, STAT5, transforming growth factor (TGF)-β and tumour necrosis factor (TNF)-α. As inflammation is a suspected suppressor of HSC function,[294,295] this suggests that in humans, leukaemic stem cells and their progeny might be creating their own self-supporting niche[296] that suppresses normal non-mutant HSCs. Moreover, response of patients to TKI treatment could be predicted by the inflammatory signalling changes observed in non-mutant HSCs – again something that could only be detected by being able to study these HSCs separately.

However, BCR-ABL, unlike other mutations, only requires a single genotyping assay to be developed for all patients, loss-of-function mutations would be less straightforward to genotype via the transcriptome in a robust and scalable manner. Efforts to scale this to more mutations have nevertheless begun in diseases such as AML[297] but the efficiencies are not yet at a point where there is a high confidence in calling the absence of a mutation. To address this issue Rodriguez-Meira et al. (2019)[298]

have developed a technique combining scRNAseq with targeted mutation sequencing from genomic (g)DNA and cDNA, allowing much more accurate calling of mutations for which cDNA expression is undetectable or highly allelic-biased.

Using this technique, the group analysed HSPCs from MPN patients with a JAK2 V617F mutation (either alone or in combination with collaborating mutations such as in TET2, EZH2, and ASXL1). The group was able to accurately determine mutation burden, order of mutation acquisition, and distinguish heterozygous and homozygous JAK2 mutations. As in CML patients, the group determined that non-mutant cells in patients are transcriptionally distinct from normal controls with patient cells having enrichment of inflammatory pathways (TNF-α, and IFN signalling), MPN microenvironments effect on WT cells from same patient.

In agreement with the mouse model data in this thesis, Rodriguez-Meira et al. reported distinct and biologically relevant molecular signatures of HSPC subclones with different mutational profiles. They also reported that cells with mutations in epigenetic regulators had a distinct transcriptional signature whereas the cells with JAK HET more closely resembled WT cells. This agrees with the data in this thesis which shows a more subtle difference between JAK2 mutant and WT cell than WT cells and cells with a TET mutation.

As the approach taken by Rodriguez-Meira et al. was a whole genome scRNAseq method, whereas in this thesis a targeted multiplex qPCR approach focusing only on stem cell and self-renewal associated genes was used, it is not surprising that Rodriguez-Meira et al. report a greater separation between mutant and WT cells, as other pathways involved in processes such as differentiation can also contribute to the transcriptome. As a consequence of scRNAseq reporting an unbiased but shallow transcriptome, the key self-renewal genes identified in this study were not identified in the patient data, this does not necessarily mean that these genes do not play a role in the self-renewal of patient MPN HSCs, but that technical limitations preclude our ability to see their expression. It may be that a future technical development will allow addition of targeted gene expression on top the techniques developed by Giustacchini

et al. and Rodriguez-Meira et al., however until then the expression of stem cell genes in human HSCs with different mutational profiles will remain somewhat unknown.

The second major development turns away from investigation into the molecular nature of malignancy, towards the monitoring and understanding of HSC clonal dynamics in humans, to understand how a single clone can emerge to drive a leukaemia. This was a project I was involved with alongside of my primary project, where we used somatic mutation acquisition as an endogenous barcode to study the relative contribution of clones over the lifetime of an individual.[299] This study was based on the fact that somatic mutations are acquired in a linear fashion throughout life, and are stably passed down, marking the clonal history of a cell.

We isolated stem and progenitor cells from the peripheral blood and BM of a healthy 59-year-old individual, and clonally grew 140 colonies which were then whole genome sequenced. Subsequently, at time points, we isolated mature cells from the same individual, which were subject to targeted bait set sequencing allowing their 'relatedness' to the original HSPCs to be determined. The whole genome sequencing of HSPC derived colonies allowed the construction of a phylogenetic tree of the haematopoietic system of this individual, and revealed that all blood cells were derived from a common ancestor that predated gastrulation (common ancestor with buckle epithelium), that in early life the stem cell population grew at a steady rate, and reached a plateau by adolescence. Based on the branching pattern of the hierarchy, computational simulations estimated of the absolute number of HSCs in this unperturbed human individual as in the range of 50,000–200,000, with an average self-renewal division rate of 2-20 months. From the follow-up time points it was observed that many HSCs were contributing to blood production and identified multilineage clones, contributing to both granulocyte and B lymphocyte lineages.

In this paper we reported that mature cell contribution was stable across the months assessed, we are continuing to collect material from this individual, this will allow investigation into the stability of clonal contribution across years and shed light of the effects of aging in the haematopoietic system. A potential future experiment is to repeat this experiment with MPN patient samples; investigating how acquisition of MPN associated mutations affects the dynamics of clonal growth, both on the mutant clone

and the surrounding non-mutant clones. This could shed light upon the relative effects of different mutations on clonal expansion, and the length of time between an individual acquiring a mutation and the presentation of disease.

At present, such techniques are not possible in large numbers of humans, but if similar approaches could be devised, studying how clonal dynamics change upon acquisition of oncogenic mutations and exposure to environmental or endogenous stresses may well become a fruitful line of enquiry. Until then, *in vitro* approaches and mouse models will be heavily relied upon. The emergence of these new single cell and clonal technologies has generated significant enthusiasm amongst researchers and clinicians trying to understand the molecular differences between malignant and non-malignant cells, including a more complete understanding of clonal competition during disease establishment, maintenance and progression. As mentioned above, the first studies are only now starting to emerge as technologies converge; the next few years promise to deliver an increasingly detailed understanding of the molecules governing fate choice in normal and malignant HSCs.

7

References

1. Kent, D. G. & Eaves, C. J. Adult Hematopoiesis. *Encyclopedia of Immunobiology* 15–25 (2016).
2. Lunger, I., Fawaz, M. & Rieger, M. A. Single-cell analyses to reveal hematopoietic stem cell fate decisions. *FEBS Lett.* **591**, 2195–2212 (2017).
3. Davila, M. L. *et al.* Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. *Sci. Transl. Med.* **6**, 224–25 (2014).
4. Jacobson, L. O., Marks, E. K., Robson, M. & Zirkle, R. E. The effect of spleen protection on mortality following x-irradiation. *J. Lab. Clin. Med.* **34**, 1538–1543 (1949).
5. Lorenz, E., Congdon, C. & Uphoff, D. Modification of acute irradiation injury in mice and guinea-pigs by bone marrow injections. *Radiology* **58**, 863–877 (1952).
6. Barnes, D. W. H. & Loutit, J. F. What is the recovery factor in spleen? *Nucleonics* **12**, 68–71 (1954).
7. Main, J. M. & Prehn, R. T. Successful skin homografts after the administration of high dosage x radiation and homologous bone marrow. *J. Natl. Cancer Inst.* **15**, 1023–1029 (1955).
8. Till, J. E. & McCulloch, E. A. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat. Res.* **14**, 213–222 (1961).
9. Becker, A. J., McCulloch, E. A. & Till, J. E. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature* **197**, 452–454 (1963).
10. Osawa, M., Hanada, K., Hamada, H. & Nakauchi, H. Long-Term Lymphohematopoietic Reconstitution by a Single CD34-Low/Negative Hematopoietic Stem Cell. *Science (80-.).* 242–245 (1996).
11. Trevisan, M., Yan, X. Q. & Iscove, N. N. Cycle initiation and colony formation in culture by murine marrow cells with long-term reconstituting potential in vivo. *Blood* **88**, 4149–4158 (1996).
12. Goodell, M. A., Brose, K., Paradis, G., Conner, A. S. & Mulligan, R. C. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *J. Exp. Med.* **183**, 1797–1806 (1996).
13. Hodgson, G. S. & Bradley, T. R. Properties of Haematopoietic stem cells surviving 5-Fluorouracil treatment: Evidence for a pre-CFU-S cell? *Nature*

381–382 (1979).
14. Ploemacher, R. E., Van Os, R., Van Beurden, C. A. J. & Down, J. D. Murine haemopoietic stem cells with long-term engraftment and marrow repopulating ability are more resistant to gamma-radiation than are spleen colony forming cells. *Int. J. Radiat. Biol.* **61**, 489–499 (1992).
15. Visser, J. W. M., Bauman, J. G. J., Mulder, A. H., Euason, J. F. & De Leeuw, A. M. Isolation of murine pluripotent hemopoietic stem cells. *J. Exp. Med.* **159**, 1576–1590 (1984).
16. Spangrude, G. J., Heimfeld, S. & Weissman, I. L. Purification and characterization of mouse hematopoietic stem cells. *Science* **241**, 58–62 (1988).
17. Okada, S. *et al.* In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. *Blood* **80**, 3044–3050 (1992).
18. Balazs, A. B., Fabian, A. J., Esmon, C. T. & Mulligan, R. C. Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. *Blood* **107**, 2317–2321 (2006).
19. Benveniste, P. *et al.* Intermediate-Term Hematopoietic Stem Cells with Extended but Time-Limited Reconstitution Potential. *Cell Stem Cell* **6**, 48–58 (2010).
20. Kiel, M. J. *et al.* SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109–1121 (2005).
21. Adolfsson, J. *et al.* Upregulation of Flt3 expression within the bone marrow Lin-Sca1+c-kit+ stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* **15**, 659–669 (2001).
22. Beerman, I. *et al.* Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc. Natl. Acad. Sci.* **107**, 5465–5470 (2010).
23. Challen, G. A., Boles, N. C., Chambers, S. M. & Goodell, M. A. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell* **6**, 265–78 (2010).
24. Kent, D. G. *et al.* Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood* **113**, 6342–50 (2009).
25. Morita, Y., Ema, H. & Nakauchi, H. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J. Exp. Med.* **207**, 1173–82 (2010).
26. Schulte, R. *et al.* Index sorting resolves heterogeneous murine hematopoietic stem cell populations. *Exp. Hematol.* **43**, 803–811 (2015).
27. Wilson, N. K. *et al.* Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* **16**, 712–724 (2015).
28. Belluschi, S. *et al.* Myelo-lymphoid lineage restriction occurs in the human haematopoietic stem cell compartment before lymphoid-primed multipotent progenitors. *Nat. Commun.* **9**, 4100 (2018).
29. Notta, F. *et al.* Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science (80-.).* **333**, 218–221 (2011).
30. Ottersbach, K. Endothelial-to-haematopoietic transition: An update on the process of making blood. *Biochem. Soc. Trans.* **47**, 591–601 (2019).
31. Mikkola, H. K. A. & Orkin, S. H. The journey of developing hematopoietic

stem cells. *Development* **133**, 3733–3744 (2006).
32. Bowie, M. B. *et al.* Identification of a new intrinsically timed developmental checkpoint that reprograms key hematopoietic stem cell properties. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 5878–5882 (2007).
33. Pinho, S. & Frenette, P. S. Haematopoietic stem cell activity and interactions with the niche. *Nat. Rev. Mol. Cell Biol.* **20**, 303–320 (2019).
34. Liu, L. & Rando, T. A. Manifestations and mechanisms of stem cell aging. *J. Cell Biol.* **193**, 257–266 (2011).
35. Morrison, S. J., Wandycz, A. M., Akashi, K., Globerson, A. & Weissman, I. L. The aging of hematopoietic stem cells. *Nat. Med.* **2**, 1011–1016 (1996).
36. Sudo, K., Ema, H., Morita, Y. & Nakauchi, H. Age-associated characteristics of murine hematopoietic stern cells. *J. Exp. Med.* **192**, 1273–1280 (2000).
37. Dykstra, B., Olthof, S., Schreuder, J., Ritsema, M. & Haan, G. De. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J. Exp. Med.* **208**, 2691–2703 (2011).
38. Cho, R. H., Sieburg, H. B. & Muller-Sieburg, C. E. A new mechanism for the aging of hematopoietic stem cells: Aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood* **111**, 5553–5561 (2008).
39. De Haan, G. & Van Zant, G. Intrinsic and extrinsic control of hemopoietic stem cell numbers: Mapping of a stem cell gene. *J. Exp. Med.* **186**, 529–536 (1997).
40. Rossi, D. J. *et al.* Cell intrinsic alterations underlie hematopoietic stem cell aging. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9194–9199 (2005).
41. Verovskaya, E. *et al.* Heterogeneity of young and aged murine hematopoietic stem cells revealed by quantitative clonal analysis using cellular barcoding. *Blood* **122**, 523–532 (2013).
42. Lyne, A.-M. *et al.* A track of the clones: new developments in cellular barcoding. *Exp. Hematol.* **68**, 15–20 (2018).
43. Watson, C. J. *et al.* The evolutionary dynamics and fitness landscape of clonal hematopoiesis. *Science (80-.).* **367**, 1449–1454 (2020).
44. Lee-Six, H. & Kent, D. G. Tracking hematopoietic stem cells and their progeny using whole-genome sequencing. *Exp. Hematol.* **83**, 12–24 (2020).
45. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
46. McKerrell, T. *et al.* Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep.* **10**, 1239–45 (2015).
47. Jaiswal, S. *et al.* Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
48. Kharchenko, P. V., Silberstein, L. & Scadden, D. T. Bayesian approach to single-cell differential expression analysis. *Nat. Methods* **11**, 740–742 (2014).
49. Tsankova, N., Renthal, W., Kumar, A. & Nestler, E. J. Epigenetic regulation in psychiatric disorders. *Nat. Rev. Neurosci.* **8**, 355–367 (2007).
50. Iwama, A. *et al.* Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi-1. *Immunity* **21**, 843–851 (2004).
51. Lessard, J. & Sauvageau, G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature* **423**, 255–260 (2003).
52. Rizo, A., Dontje, B., Vellenga, E., De Haan, G. & Sehuringa, J. J. Long-term maintenance of human hematopoietic stem/progenitor cells by expression of BMI1. *Blood* **111**, 2621–2630 (2008).

53. Xu, B. *et al.* The Chromatin Remodeler BPTF Activates a Stemness Gene-Expression Program Essential for the Maintenance of Adult Hematopoietic Stem Cells. *Stem Cell Reports* **10**, 675–683 (2018).
54. Roussy, M. *et al.* NUP98-BPTF gene fusion identified in primary refractory acute megakaryoblastic leukemia of infancy. *Genes Chromosom. Cancer* **57**, 311–319 (2018).
55. Hamey, F. K. *et al.* Reconstructing blood stem cell regulatory network models from single-cell molecular profiles. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5822–5829 (2017).
56. Steinauer, N. *et al.* ETO2 Regulates Cell-Fate Genes and Controls Relapse in Acute Myeloid Leukemia. *Blood* **130**, 3808–3808 (2017).
57. Jung, J. *et al.* CBX7 Induces Self-Renewal of Human Normal and Malignant Hematopoietic Stem and Progenitor Cells by Canonical and Non-canonical Interactions. *Cell Rep.* **26**, 1906–1918 (2019).
58. Stepanova, L. & Sorrentino, B. P. A limited role for p16Ink4a and p19Arf in the loss of hematopoietic stem cells during proliferative stress. *Blood* **106**, 827–832 (2005).
59. Akala, O. O. *et al.* Long-term haematopoietic reconstitution by Trp53-/-p16 Ink4a-/-p19Arf-/- multipotent progenitors. *Nature* **453**, 228–232 (2008).
60. Zhao, R., Choi, B. Y., Lee, M. H., Bode, A. M. & Dong, Z. Implications of Genetic and Epigenetic Alterations of CDKN2A (p16INK4a) in Cancer. *EBioMedicine* **8**, 30–39 (2016).
61. Stanley, E. R. & Chitu, V. CSF-1 receptor signaling in myeloid cells. *Cold Spring Harb. Perspect. Biol.* **6**, (2014).
62. Mossadegh-Keller, N. *et al.* M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature* **497**, 239–243 (2013).
63. Edwards, D. K. *et al.* CSF1R inhibitors exhibit antitumor activity in acute myeloid leukemia by blocking paracrine signals from support cells. *Blood* **133**, 588–599 (2019).
64. Jeong, M. *et al.* Loss of Dnmt3a Immortalizes Hematopoietic Stem Cells In Vivo Data and Software Availability GSE98191 Loss of Dnmt3a Immortalizes Hematopoietic Stem Cells In Vivo. *Cell Rep.* **23**, 1–10 (2018).
65. Challen, G. A. *et al.* Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* **44**, 23–31 (2012).
66. Nishida, C. *et al.* Epidermal Growth Factor-like Domain 7 Promotes Hematopoietic Stem Cell Expansion and Increases Myeloid-Megakaryocytic Lineage Priming through beta3 Integrin. *Blood* **124**, 2919–2919 (2014).
67. Papaioannou, D. *et al.* Prognostic and biological significance of the proangiogenic factor EGFL7 in acute myeloid leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E4641–E4647 (2017).
68. Knudsen, K. J. *et al.* ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation. *Genes Dev.* **29**, 1915–1929 (2015).
69. Kamminga, L. M. *et al.* The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* **107**, 2170–2179 (2006).
70. Kruse, E. A. *et al.* Dual requirement for the ETS transcription factors Fli-1 and Erg in hematopoietic stem cells and the megakaryocyte lineage. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13814–13819 (2009).
71. Masuya, M. *et al.* Dysregulation of granulocyte, erythrocyte, and NK cell lineages in Fli-1 gene-targeted mice. *Blood* **105**, 95–102 (2005).
72. Kornblau, S. M. *et al.* Abnormal expression of FLI1 protein is an adverse

prognostic factor in acute myeloid leukemia. *Blood* **118**, 5604–5612 (2011).
73. Miyamoto, K. *et al.* Foxo3a Is Essential for Maintenance of the Hematopoietic Stem Cell Pool. *Cell Stem Cell* **1**, 101–112 (2007).
74. Santamaría, C. M. *et al.* High FOXO3a expression is associated with a poorer prognosis in AML with normal cytogenetics. *Leuk. Res.* **33**, 1706–1709 (2009).
75. Takai, J. *et al.* The Gata1 5′ region harbors distinct cis-regulatory modules that direct gene activation in erythroid cells and gene inactivation in HSCs. *Blood* **122**, 3450–3460 (2013).
76. Crispino, J. D. & Horwitz, M. S. GATA factor mutations in hematologic disease. *Blood* **129**, 2103–2110 (2017).
77. Rodrigues, N. P., Tipping, A. J., Wang, Z. & Enver, T. GATA-2 mediated regulation of normal hematopoietic stem/progenitor cell function, myelodysplasia and myeloid leukemia. *Int. J. Biochem. Cell Biol.* **44**, 457–460 (2012).
78. Yokota, T. & Kanakura, Y. Genetic abnormalities associated with acute lymphoblastic leukemia. *Cancer Sci.* **107**, 721–725 (2016).
79. Ku, C. J., Hosoya, T., Maillard, I. & Engel, J. D. GATA-3 regulates hematopoietic stem cell maintenance and cell-cycle entry. *Blood* **119**, 2242–2251 (2012).
80. Zeng, H., Yücel, R., Kosan, C., Klein-Hitpass, L. & Möröy, T. Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. *EMBO J.* **23**, 4116–4125 (2004).
81. Van Der Meer, L. T., Jansen, J. H. & Van Der Reijden, B. A. Gfi1 and Gfi1b: Key regulators of hematopoiesis. *Leukemia* **24**, 1834–1843 (2010).
82. Hönes, J. M. *et al.* GFI1 as a novel prognostic and therapeutic factor for AML/MDS. *Leukemia* **30**, 1237–1245 (2016).
83. Khandanpour, C. *et al.* Evidence that growth factor independence 1b regulates dormancy and peripheral blood mobilization of hematopoietic stem cells. *Blood* **116**, 5149–5161 (2010).
84. Anguita, E., Candel, F. J., Chaparro, A. & Roldán-Etcheverry, J. J. Transcription factor GFI1B in health and disease. *Front. Oncol.* **7**, 28 (2017).
85. Jackson, J. T. *et al.* Hhex Regulates Hematopoietic Stem Cell Self-Renewal and Stress Hematopoiesis via Repression of Cdkn2a. *Stem Cells* **35**, 1948–1957 (2017).
86. Shields, B. J. *et al.* Acute myeloid leukemia requires Hhex to enable PRC2-mediated epigenetic repression of Cdkn2a. *Genes Dev.* **30**, 78–91 (2016).
87. Yang, D. *et al.* Enforced expression of Hoxa5 in haematopoietic stem cells leads to aberrant erythropoiesis in vivo. *Cell Cycle* **14**, 612–620 (2015).
88. Collins, C. T. & Hess, J. L. Role of HOXA9 in leukemia: Dysregulation, cofactors and essential targets. *Oncogene* **35**, 1090–1098 (2016).
89. Umeda, S. *et al.* Prognostic significance of HOXB 4 in de novo acute myeloid leukemia. *Hematology* **17**, 125–131 (2012).
90. Antonchuk, J., Sauvageau, G. & Humphries, R. K. HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. *Cell* **109**, 39–45 (2002).
91. Li, Z. *et al.* Leukaemic alterations of IKZF1 prime stemness and malignancy programs in human lymphocytes. *Cell Death Dis.* **9**, 1–11 (2018).
92. Bernitz, J. M., Kim, H. S., MacArthur, B., Sieburg, H. & Moore, K. Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions. *Cell* **167**, 1296–1309 (2016).
93. Umemoto, T. *et al.* Integrin-αvβ3 regulates thrombopoietin-mediated

maintenance of hematopoietic stem cells. *Blood* **119**, 83–94 (2012).
94. Nurden, A. T. & Pillois, X. ITGA2B and ITGB3 gene mutations associated with Glanzmann thrombasthenia. *Platelets* **29**, 98–101 (2018).
95. Thorén, L. A. *et al.* Kit Regulates Maintenance of Quiescent Hematopoietic Stem Cells. *J. Immunol.* **180**, 2045–2053 (2008).
96. Ma, Y. *et al.* The c-KIT mutation causing human mastocytosis is resistant to STI571 and other KIT kinase inhibitors; kinases with enzymatic site mutations show different inhibitor sensitivity profiles than wild-type kinases and those with regulatory-type mutations. *Blood* **99**, 1741–1744 (2002).
97. Cleveland, S. M. *et al.* Lmo2 induces hematopoietic stem cell-like features in T-cell progenitor cells prior to leukemia. *Stem Cells* **31**, 882–894 (2013).
98. Souroullas, G. P., Salmon, J. M., Sablitzky, F., Curtis, D. J. & Goodell, M. A. Adult Hematopoietic Stem and Progenitor Cells Require Either Lyl1 or Scl for Survival. *Cell Stem Cell* **4**, 180–186 (2009).
99. Zhang, Y. *et al.* PR-domain - containing Mds1-Evi1 is critical for long-term hematopoietic stem cell function. *Blood* **118**, 3853–3861 (2011).
100. Alpermann, T. *et al.* Prognosis of Mecom(EVI1)- rearranged MDS and AML Patients Strongly Depends on Accompanying Molecular Mutations but Not on Blast Counts. *Blood* **126**, 1372–1372 (2015).
101. Unnisa, Z. *et al.* Meis1 preserves hematopoietic stem cells in mice by limiting oxidative stress. *Blood* **120**, 4973–4981 (2012).
102. Dickson, G. J. *et al.* HOXA/PBX3 knockdown impairs growth and sensitizes cytogenetically normal acute myeloid leukemia cells to chemotherapy. *Haematologica* **98**, 1216–1225 (2013).
103. Liu, J. *et al.* Meis1 is critical to the maintenance of human acute myeloid leukemia cells independent of MLL rearrangements. *Ann. Hematol.* **96**, 567–574 (2017).
104. Lee, Y. N. *et al.* KIT signaling regulates MITF expression through miRNAs in normal and malignant mast cell proliferation. *Blood* **117**, 3629–3640 (2011).
105. Li, Y. *et al.* The transcription factors GATA2 and microphthalmia-associated transcription factor regulate Hdc gene expression in mast cells and are required for IgE/mast cell–mediated anaphylaxis. *J. Allergy Clin. Immunol.* **142**, 1173–1184 (2018).
106. Petit-Cocault, L., Volle-Challier, C., Fleury, M., Péault, B. & Souyri, M. Dual role of Mpl receptor during the establishment of definitive hematopoiesis. *Development* **134**, 3031–3040 (2007).
107. Pattabiraman, D. R. & Gonda, T. J. Role and potential for therapeutic targeting of MYB in leukemia. *Leukemia* **27**, 269–277 (2013).
108. Lieu, Y. K. & Reddy, E. P. Conditional c-myb knockout in adult hematopoietic stem cells leads to loss of self-renewal due to impaired proliferation and accelerated differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21689–21694 (2009).
109. Di Tullio, A., Passaro, D., Rouault-Pierre, K., Purewal, S. & Bonnet, D. Nuclear Factor Erythroid 2 Regulates Human HSC Self-Renewal and T Cell Differentiation by Preventing NOTCH1 Activation. *Stem Cell Reports* **9**, 5–11 (2017).
110. Jutzi, J. S. *et al.* MPN patients harbor recurrent truncating mutations in transcription factor NF-E2. *J. Exp. Med.* **210**, 1003–1019 (2013).
111. Ficara, F., Murphy, M. J., Lin, M. & Cleary, M. L. Pbx1 Regulates Self-Renewal of Long-Term Hematopoietic Stem Cells by Maintaining Their

Quiescence. *Cell Stem Cell* **2**, 484–496 (2008).
112. Aguilo, F. *et al.* Prdm16 is a physiologic regulator of hematopoietic stem cells. *Blood* **117**, 5057–5066 (2011).
113. Yu, H. *et al.* Downregulation of Prdm16 mRNA is a specific antileukemic mechanism during HOXB4-mediated HSC expansion in vivo. *Blood* **124**, 1737–1747 (2014).
114. Gur-Cohen, S. *et al.* EPCR Guides Hematopoietic Stem Cells Homing to the Bone Marrow Independently of Niche Clearance. *Blood* **128**, 4538–4538 (2016).
115. Cai, X. *et al.* Runx1 loss minimally impacts long-term hematopoietic stem cells. *PLoS One* **6**, (2011).
116. Licht, J. D. AML1 and the AML1-ETO fusion protein in the pathogenesis of t(8;21) AML. *Oncogene* **20**, 5660–5679 (2001).
117. Kuo, M. C. *et al.* RUNX1 mutations are frequent in chronic myelomonocytic leukemia and mutations at the C-terminal region might predict acute myeloid leukemia transformation. *Leukemia* **23**, 1426–1431 (2009).
118. Dempsey, L. A. Pu.1 in HSCs. *Nat. Immunol.* **14**, 427 (2013).
119. Antony-Debré, I. *et al.* Pharmacological inhibition of the transcription factor PU.1 in leukemia. *J. Clin. Invest.* **127**, 4297–4313 (2017).
120. Maslah, N., Cassinat, B., Verger, E., Kiladjian, J. J. & Velazquez, L. The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. *Leukemia* **31**, 1661–1670 (2017).
121. Ema, H. *et al.* Quantification of self-renewal capacity in single hematopoietic stem cells from normal and Lnk-deficient mice. *Dev. Cell* **8**, 907–914 (2005).
122. Bersenev, A., Wu, C., Balcerek, J. & Tong, W. Lnk controls mouse hematopoietic stem cell self-renewal and quiescence through direct interactions with JAK2. *J. Clin. Invest.* **118**, 2832–2844 (2008).
123. Prasad, P., Lennartsson, A. & Ekwall, K. The roles of SNF2/SWI2 nucleosome remodeling enzymes in blood cell differentiation and leukemia. *Biomed Res. Int.* **2015**, (2015).
124. Deneault, E. *et al.* A Functional Screen to Identify Novel Effectors of Hematopoietic Stem Cell Activity. *Cell* **137**, 369–379 (2009).
125. Sanda, T. & Leong, W. Z. TAL1 as a master oncogenic transcription factor in T-cell acute lymphoblastic leukemia. *Exp. Hematol.* **53**, 7–15 (2017).
126. Wu, J. Q. *et al.* Tcf7 Is an Important Regulator of the Switch of Self-Renewal and Differentiation in a Multipotential Hematopoietic Cell Line. *PLoS Genet.* **8**, e1002565 (2012).
127. Yu, S. *et al.* Hematopoietic and leukemic stem cells have distinct dependence on Tcf1 and Lef1 transcription factors. *J. Biol. Chem.* **291**, 11148–11160 (2016).
128. Solary, E., Bernard, O. A., Tefferi, A., Fuks, F. & Vainchenker, W. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. *Leukemia* **28**, 485–496 (2014).
129. Ohoka, N., Yoshii, S., Hattori, T., Onozaki, K. & Hayashi, H. TRB3, a novel ER stress-inducible gene, is induced via ATF4-CHOP pathway and is involved in cell death. *EMBO J.* **24**, 1243–1255 (2005).
130. Salomé, M., Hopcroft, L. & Keeshan, K. Inverse and correlative relationships between TRIBBLES genes indicate non-redundant functions during normal and malignant hemopoiesis. *Exp. Hematol.* **66**, 63–78 (2018).
131. Sigurdsson, V. *et al.* CD244 Marks Non-Functional Hematopoietic Stem Cells

with a Mast Cell Signature after Induction of Endoplasmic Reticulum Stress. *Blood* **134**, 2474–2474 (2019).
132. Sanjuan-Pla, A. *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502**, 232–236 (2013).
133. Bradley, T. R. & Metcalf, D. The growth of mouse bone marrow cells in vitro. *Aust. J. Exp. Biol. Med. Sci.* **44**, 287–299 (1966).
134. Holmes, R. & Zuñiga-Pflücker, J. C. The OP9-DL1 system: Generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro. *Cold Spring Harb. Protoc.* **4**, (2009).
135. Moran-Crusio, K. *et al.* Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-Renewal and Myeloid Transformation. *Cancer Cell* **20**, 11–24 (2011).
136. Guryanova, O. A. *et al.* Dnmt3a regulates myeloproliferation and liver-specific expansion of hematopoietic stem and progenitor cells. *Leukemia* **30**, 1133–1142 (2016).
137. de Haan, G., Nijhof, W. & Van Zant, G. Mouse Strain-Dependent Changes in Frequency and Proliferation of Hematopoietic Stem Cells During Aging: Correlation Between Lifespan and Cycling Activity. *Blood* **89**, 1543–1550 (1997).
138. Purton, L. E. & Scadden, D. T. Limiting Factors in Murine Hematopoietic Stem Cell Assays. *Cell Stem Cell* **1**, 263–270 (2007).
139. Lu, R., Neff, N. F., Quake, S. R. & Weissman, I. L. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nat. Biotechnol.* **29**, 928–934 (2011).
140. Pei, W. *et al.* Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* **548**, 456–460 (2017).
141. Santaguida, M. *et al.* JunB Protects against Myeloid Malignancies by Limiting Hematopoietic Stem Cell Proliferation and Differentiation without Affecting Self-Renewal. *Cancer Cell* **15**, 341–352 (2009).
142. Lundberg, P. *et al.* Myeloproliferative neoplasms can be initiated from a single hematopoietic stem cell expressing JAK2-V617F. *J. Exp. Med.* **211**, 2213–2230 (2014).
143. Lemischka, I. R., Raulet, D. H. & Mulligan, R. C. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* **45**, 917–27 (1986).
144. Jordan, C. T. & Lemischka, I. R. Clonal and systemic analysis of long-term hematopoiesis in the mouse. *Genes Dev.* **4**, 220–32 (1990).
145. Dick, J. E., Magli, M. C., Huszar, D., Phillips, R. A. & Bernstein, A. Introduction of a Selectable Gene into Primitive Stem Cells Capable of Long-Term Reconstitution of the Hemopoietic System of W/W v Mice. *Cell* **42**, 71–79 (1985).
146. Keller, G. & Snodgrass, R. Life span of multipotential hematopoietic stem cells in vivo. *J. Exp. Med.* **171**, 1407–18 (1990).
147. Naik, S. H. *et al.* Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229–232 (2013).
148. Rodriguez-Fraticelli, A. E. *et al.* Clonal analysis of lineage fate in native haematopoiesis. *Nature* **553**, 212–216 (2018).
149. Paul, F. *et al.* Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* **163**, 1663–1677 (2015).
150. Olsson, A. *et al.* Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* **537**, 698–702 (2016).
151. Laurenti, E. & Göttgens, B. From haematopoietic stem cells to complex

differentiation landscapes. *Nature* **553**, 418–426 (2018).
152. Wilson, N. K. & Göttgens, B. Single-Cell Sequencing in Normal and Malignant Hematopoiesis. *Hemasphere* **2**, (2018).
153. Fialkow, P. J. *et al.* Acute Nonlymphocytic Leukemia: Expression in Cells Restricted to Granulocytic and Monocytic Differentiation. *N. Engl. J. Med.* **301**, 1–5 (1979).
154. Nowell, P. C. & Hungerford, D. A. A minute chromosome in human chronic granulocytic leukemia. *Science (80-.).* **132**, 1497 (1960).
155. Hochhaus, A. *et al.* Chronic myeloid leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* Supplement 4: iv41-iv51 (2017).
156. Kantarjian, H. *et al.* Improved survival in chronic myeloid leukemia since the introduction of imatinib therapy: A single-institution historical experience. *Blood* **119**, 1981–1987 (2012).
157. Tefferi, A. & Pardanani, A. Myeloproliferative neoplasms: A contemporary review. *JAMA Oncol.* **1**, 97–105 (2015).
158. Moulard, O. *et al.* Epidemiology of myelofibrosis, essential thrombocythemia, and polycythemia vera in the European Union. *Eur. J. Haematol.* **92**, 289–297 (2014).
159. Barbui, T. *et al.* Polycythemia vera: The natural history of 1213 patients followed for 20 years. *Ann. Intern. Med.* **123**, 656–664 (1995).
160. Tefferi, A. *et al.* Long-term survival and blast transformation in molecularly annotated essential thrombocythemia, polycythemia vera, and myelofibrosis. *Blood* **124**, 2507–2513 (2014).
161. Marchioli, R. *et al.* Cardiovascular Events and Intensity of Treatment in Polycythemia Vera. *N. Engl. J. Med.* **368**, 22–33 (2013).
162. Tefferi, A. Polycythemia vera: A comprehensive review and clinical recommendations. *Mayo Clin. Proc.* **78**, 174–194 (2003).
163. McMullin, M. F. *et al.* A guideline for the diagnosis and management of polycythaemia vera. A British Society for Haematology Guideline. *Br. J. Haematol.* **184**, 176–191 (2019).
164. Mullally, A. *et al.* Depletion of Jak2V617F myeloproliferative neoplasm-propagating stem cells by interferon-α in a murine model of polycythemia vera. *Blood* **121**, 3692–3702 (2013).
165. Briére, J. B. Essential thrombocythemia. *Orphanet J. Rare Dis.* **2**, 3 (2007).
166. Harrison, C. N. *et al.* Guideline for investigation and management of adults and children presenting with a thrombocytosis. *Br. J. Haematol.* **149**, 352–375 (2010).
167. Barbui, T. *et al.* The 2016 WHO classification and diagnostic criteria for myeloproliferative neoplasms: document summary and in-depth discussion. *Blood Cancer J.* **8**, 15 (2018).
168. Grove, C. S. & Vassiliou, G. S. Acute myeloid leukaemia: A paradigm for the clonal evolution of cancer? *Dis. Model. Mech.* **7**, 941–951 (2014).
169. Kim, T. H. *et al.* Spectrum of somatic mutation dynamics in chronic myeloid leukemia following tyrosine kinase inhibitor therapy. *Blood* **129**, 38–47 (2017).
170. Petzer, A. L., Eaves, C. J., Barnett, M. J. & Eaves, A. C. Selective expansion of primitive normal hematopoietic cells in cytokine- supplemented cultures of purified cells from patients with chronic myeloid leukemia. *Blood* **90**, 64–69 (1997).
171. Song, J., Mercer, D., Hu, X., Liu, H. & Li, M. M. Common leukemia- and

lymphoma-associated genetic aberrations in healthy individuals. *J. Mol. Diagnostics* **13**, 213–219 (2011).
172. Ismail, S. I., Naffa, R. G., Yousef, A. M. F. & Ghanim, M. T. Incidence of bcr-abl fusion transcripts in healthy individuals. *Mol. Med. Rep.* **9**, 1271–1276 (2014).
173. Koschmieder, S. *et al.* Inducible chronic phase of myeloid leukemia with expansion of hematopoietic stem cells in a transgenic model of BCR-ABL leukemogenesis. *Blood* **105**, 324–334 (2005).
174. Schemionek, M. *et al.* BCR-ABL enhances differentiation of long-term repopulating hematopoietic stem cells. *Blood* **115**, 3185–3195 (2010).
175. Forbes, S. A. *et al.* COSMIC: Mining complete cancer genomes in the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* **39**, (2011).
176. Patel, K. P. *et al.* Correlation of mutation profile and response in patients with myelofibrosis treated with ruxolitinib. *Blood* **126**, 790–797 (2015).
177. Li, N. *et al.* Frequency and allele burden of CALR mutations in Chinese with essential thrombocythemia and primary myelofibrosis without JAK2V617F or MPL mutations. *Leuk. Res.* **39**, 510–514 (2015).
178. Wilmes, S. *et al. Mechanism of homodimeric cytokine receptor activation and dysregulation by oncogenic mutations*. http://science.sciencemag.org/.
179. Chen, E., Staudt, L. M. & Green, A. R. Janus Kinase Deregulation in Leukemia and Lymphoma. *Immunity* **36**, 529–541 (2012).
180. Baxter, E. J. *et al.* Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* **365**, 1054–1061 (2005).
181. James, C. *et al.* A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* **434**, 1144–1148 (2005).
182. Jones, A. V. *et al.* Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. *Blood* **106**, 2162–2168 (2005).
183. Kralovics, R. *et al.* Altered gene expression in myeloproliferative disorders correlates with activation of signaling by the V617F mutation of Jak2. *Blood* **106**, 3374–3376 (2005).
184. Levine, R. L. *et al.* Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell* **7**, 387–397 (2005).
185. Scott, L. M. *et al.* The V617F JAK2 mutation is uncommon in cancers and in myeloid malignancies other than the classic myeloproliferative disorders. *Blood* **106**, 2920–2921 (2005).
186. Jelinek, J. *et al.* JAK2 mutation 1849G>T is rare in acute leukemias but can be found in CMML, Philadelphia chromosome-negative CML, and megakaryocytic leukemia. *Blood* **106**, 3370–3373 (2005).
187. Zhao, R. *et al.* Identification of an acquired JAK2 mutation in polycythemia vera. *J. Biol. Chem.* **280**, 22788–22792 (2005).
188. Passamonti, F. & Rumi, E. Clinical relevance of JAK2 (V617F) mutant allele burden. *Haematologica* **94**, 7–10 (2009).
189. Campbell, P. J. *et al.* Definition of subtypes of essential thrombocythaemia and relation to polycythaemia vera based on JAK2 V617F mutation status: A prospective study. *Lancet* **366**, 1945–1953 (2005).
190. Ortmann, C. A. *et al.* Effect of Mutation Order on Myeloproliferative Neoplasms. *N. Engl. J. Med.* **372**, 601–612 (2015).
191. Li, J., Kent, D. G., Chen, E. & Green, A. R. Mouse models of myeloproliferative neoplasms: JAK of all grades. *DMM Dis. Model. Mech.* **4**,

311–317 (2011).
192. Xu, X. *et al.* JAK2V617F: Prevalence in a large Chinese hospital population. *Blood* **109**, 339–342 (2007).
193. Campbell, P. J. Somatic and germline genetics at the JAK2 locus. *Nat. Methods* **41**, 385–386 (2009).
194. Chou, F.-S. & Mulloy, J. C. The Thrombopoietin/MPL pathway in hematopoiesis and leukemogenesis. *J. Cell. Biochem.* **112**, 1491–1498 (2011).
195. Sangkhae, V., Etheridge, S. L., Kaushansky, K. & Hitchcock, I. S. The thrombopoietin receptor, MPL, is critical for development of a JAK2V617F-induced myeloproliferative neoplasm. *Blood* **124**, 3956–3963 (2014).
196. Defour, J. P. *et al.* Tryptophan at the transmembrane-cytosolic junction modulates thrombopoietin receptor dimerization and activation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2540–2545 (2013).
197. Vainchenker, W. & Kralovics, R. Genetic basis and molecular pathophysiology of classical myeloproliferative neoplasms. *Blood* **129**, 667–679 (2017).
198. Pikman, Y. *et al.* MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med.* **3**, 1140–1151 (2006).
199. Michalak, M., Corbett, E. F., Mesaeli, N., Nakamura, K. & Opas, M. Calreticulin: One protein, one gene, many functions. *Biochem. J.* **344**, 281–292 (1999).
200. Araki, M. & Komatsu, N. Novel molecular mechanism of cellular transformation by a mutant molecular chaperone in myeloproliferative neoplasms. *Cancer Sci.* **108**, 1907–1912 (2017).
201. Li, J. *et al.* Mutant calreticulin knockin mice develop thrombocytosis andmyelofibrosis without a stemcell self-renewal advantage. *Blood* **131**, 649–661 (2018).
202. Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science (80-.).* **357**, (2017).
203. Shen, H. & Laird, P. W. Interplay between the cancer genome and epigenome. *Cell* vol. 153 38–55 (2013).
204. Lundberg, P. *et al.* Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. *Blood* **123**, 2220–2228 (2014).
205. Jin, B., Li, Y. & Robertson, K. D. DNA methylation: Superior or subordinate in the epigenetic hierarchy? *Genes and Cancer* **2**, 607–617 (2011).
206. Yang, L., Rau, R. & Goodell, M. A. DNMT3A in haematological malignancies. *Nat. Rev. Cancer* **15**, 152–165 (2015).
207. Emperle, M. *et al.* Mutations of R882 change flanking sequence preferences of the DNA methyltransferase DNMT3A and cellular methylation patterns. *Nucleic Acids Res.* **47**, 11355–11367 (2019).
208. Dai, Y. J. *et al.* Conditional knockin of Dnmt3a R878H initiates acute myeloid leukemia with mTOR pathway involvement. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5237–5242 (2017).
209. Saint-Martin, C. *et al.* Analysis of the Ten-Eleven Translocation 2 (TET2) gene in familial myeloproliferative neoplasms. *Blood* **114**, 1628–1632 (2009).
210. Ko, M. *et al.* Ten-eleven-translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14566–14571 (2011).
211. Li, Z. *et al.* Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**, 4509–4518 (2011).

212. Quivoron, C. *et al.* TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis. *Cancer Cell* **20**, 25–38 (2011).
213. Medeiros, B. C. *et al.* Isocitrate dehydrogenase mutations in myeloid malignancies. *Leukemia* **31**, 272–281 (2017).
214. Inoue, S. *et al.* Mutant IDH1 Downregulates ATM and Alters DNA Repair and Sensitivity to DNA Damage Independent of TET2. *Cancer Cell* **30**, 337–348 (2016).
215. Chen, F. *et al.* Oncometabolites d- and l-2-Hydroxyglutarate Inhibit the AlkB Family DNA Repair Enzymes under Physiological Conditions. *Chem. Res. Toxicol.* **30**, 1102–1110 (2017).
216. Gan, L. *et al.* Epigenetic regulation of cancer progression by EZH2: From biological insights to therapeutic potential. *Biomark. Res.* **6**, 10 (2018).
217. Guglielmelli, P. *et al.* EZH2 mutational status predicts poor survival in myelofibrosis. *Blood* **118**, 5227–5234 (2011).
218. Gelsi-Boyer, V. *et al.* Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. *J. Hematol. Oncol.* **5**, 12 (2012).
219. Carbuccia, N. *et al.* Mutations of ASXL1 gene in myeloproliferative neoplasms. *Leukemia* **23**, 2183–2186 (2009).
220. Dawson, M. A. *et al.* JAK2 phosphorylates histone H3Y41 and excludes HP1α from chromatin. *Nature* **461**, 819–822 (2009).
221. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029–1041 (2017).
222. Hoermann, G., Greiner, G. & Valent, P. Cytokine Regulation of Microenvironmental Cells in Myeloproliferative Neoplasms. *Mediators Inflamm.* **2015**, (2015).
223. Mullally, A. *et al.* Physiological Jak2V617F Expression Causes a Lethal Myeloproliferative Neoplasm with Differential Effects on Hematopoietic Stem and Progenitor Cells. *Cancer Cell* **17**, 584–596 (2010).
224. Mullally, A. *et al.* Distinct roles for long-term hematopoietic stem cells and erythroid precursor cells in a murine model of Jak2V617F-mediated polycythemia vera. *Blood* **120**, 166–172 (2012).
225. Chen, E. *et al.* Distinct effects of concomitant Jak2V617F expression and Tet2 loss in mice promote disease progression in myeloproliferative neoplasms. *Blood* **125**, 327–335 (2015).
226. Hasan, S. *et al.* JAK2V617F expression in mice amplifies early hematopoietic cells and gives them a competitive advantage that is hampered by IFNa. *Blood* **122**, 1464–1477 (2013).
227. Li, J. *et al.* JAK2 V617F impairs hematopoietic stem cell function in a conditional knock-in mouse model of JAK2 V617F-positive essential thrombocythemia. *Blood* **116**, 1528–1538 (2010).
228. Li, J. *et al.* JAK2V617F homozygosity drives a phenotypic switch in myeloproliferative neoplasms, but is insufficient to sustain disease. *Blood* **123**, 3139–3151 (2014).
229. Shepherd, M. S. *et al.* Single-cell approaches identify the molecular network driving malignant hematopoietic stem cell self-renewal. *Blood* **132**, (2018).
230. Kent, D. G. *et al.* Self-Renewal of Single Mouse Hematopoietic Stem Cells Is Reduced by JAK2V617F Without Compromising Progenitor Cell Expansion. *PLoS Biol.* **11**, (2013).

231. Kameda, T. *et al.* Loss of TET2 has dual roles in murine myeloproliferative neoplasms: Disease sustainer and disease accelerator. *Blood* **125**, 304–315 (2015).
232. Shide, K. *et al.* Development of ET, primary myelofibrosis and PV in mice expressing JAK2 V617F. *Leukemia* **22**, 87–95 (2008).
233. Shide, K. *et al.* TET2 is essential for survival and hematopoietic stem cell homeostasis. *Leukemia* **26**, 2216–2223 (2012).
234. Yang, Y., Akada, H., Nath, D., Hutchison, R. E. & Mohi, G. Loss of Ezh2 cooperates with Jak2V617F in the development of myelofibrosis in a mouse model of myeloproliferative neoplasm. *Blood* **127**, 3410–3423 (2016).
235. McKenney, A. S. *et al.* JAK2/IDH-mutant–driven myeloproliferative neoplasm is sensitive to combined targeted inhibition. *J. Clin. Invest.* **128**, 789–804 (2018).
236. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–74 (2013).
237. Loberg, M. A. *et al.* Sequentially inducible mouse models reveal that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. *Leukemia* 1635–1649 (2019) doi:10.1038/s41375-018-0368-6.
238. Poitras, J. L. *et al.* Dnmt3a deletion cooperates with the Flt3/ITD mutation to drive leukemogenesis in a murine model. *Oncotarget* **7**, 69124–69135 (2016).
239. Chan, W.-I. *et al.* The Transcriptional Coactivator Cbp Regulates Self-Renewal and Differentiation in Adult Hematopoietic Stem Cells. *Mol. Cell. Biol.* **31**, 5046–5060 (2011).
240. Vassiliou, G. S. *et al.* Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nat. Genet.* **43**, 470–476 (2011).
241. Donehower, L. A. *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215–221 (1992).
242. Kent, D. G., Dykstra, B. J. & Eaves, C. J. Isolation and Assessment of Single Long-Term Reconstituting Hematopoietic Stem Cells from Adult Mouse Bone Marrow. *Curr. Protoc. Stem Cell Biol.* **38**, 4.1-4.24 (2016).
243. Dykstra, B. *et al.* Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell* **1**, 218–29 (2007).
244. Moignard, V. *et al.* Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat. Cell Biol.* **15**, 363–372 (2013).
245. Guo, G. *et al.* Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Dev. Cell* **18**, 675–685 (2010).
246. Pernod, G., Fish, R., Liu, J. W. & Kruithof, E. K. O. Increasing lentiviral vector titer using inhibitors of protein kinase R. *Biotechniques* **36**, 576–580 (2004).
247. Pfaffl, M. W. Quantification strategies in real-time PCR. in *A-Z of quantitative PCR* (ed. Bustin, S. A.) 87–112 (International Univerity Line, 2004).
248. Kent, D. G., Dykstra, B. J., Cheyne, J., Elaine, M. & Eaves, C. J. Steel factor coordinately regulates the molecular signature and biologic function of hematopoietic stem cells. *Blood* **112**, 560–567 (2008).
249. Mullighan, C. G. *et al.* CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* **471**, 235–241 (2011).
250. Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471**, 189–196 (2011).

251. Kung, A. L. *et al.* Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev.* **14**, 272–277 (2000).
252. Kang-Decker, N. *et al.* Loss of CBP causes T cell lymphomagenesis in synergy with p27 Kip1 insufficiency. *Cancer Cell* **5**, 177–189 (2004).
253. Falini, B. *et al.* Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N. Engl. J. Med.* **352**, 254–266 (2005).
254. Grimwade, D., Ivey, A. & Huntly, B. J. P. Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. *Blood* **127**, 29–41 (2016).
255. Falini, B. *et al.* Altered nucleophosmin transport in acute myeloid leukaemia with mutated NPM1: Molecular basis and clinical implications. *Leukemia* **23**, 1731–1743 (2009).
256. Brunetti, L. *et al.* Mutant NPM1 Maintains the Leukemic State through HOX Expression. *Cancer Cell* **34**, 499–512 (2018).
257. Levine, A. J. & Oren, M. The first 30 years of p53: Growing ever more complex. *Nat. Rev. Cancer* **9**, 749–758 (2009).
258. Chen, S. *et al.* Mutant p53 drives clonal hematopoiesis through modulating epigenetic pathway. *Nat. Commun.* **10**, 1–14 (2019).
259. Xie, M. *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).
260. TeKippe, M., Harrison, D. E. & Chen, J. Expansion of hematopoietic stem cell phenotype and activity in Trp53-null mice. *Exp. Hematol.* **31**, 521–527 (2003).
261. Liu, Y. *et al.* p53 Regulates Hematopoietic Stem Cell Quiescence. *Cell Stem Cell* **4**, 37–48 (2009).
262. Benz, C. *et al.* Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell Stem Cell* **10**, 273–283 (2012).
263. Chen, Z. *et al.* Inference of immune cell composition on the expression profiles of mouse tissue. *Sci. Rep.* **7**, 1–11 (2017).
264. Koldej, R. M. *et al.* Comparison of insulators and promoters for expression of the Wiskott-Aldrich syndrome protein using lentiviral vectors. *Hum. Gene Ther. Clin. Dev.* **24**, 77–85 (2013).
265. Song, M. K., Park, B. B. & Uhm, J. E. Understanding splenomegaly in myelofibrosis: Association with molecular pathogenesis. *Int. J. Mol. Sci.* **19**, (2018).
266. Accurso, V. *et al.* Splenomegaly impacts prognosis in essential thrombocythemia and polycythemia vera: A single center study. *Hematol. Rep.* **11**, 95–97 (2019).
267. Bhatia, M., Wang, J. C. Y., Kapp, U., Bonnet, D. & Dick, J. E. Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5320–5325 (1997).
268. Velasco-Hernandez, T., Säwén, P., Bryder, D. & Cammenga, J. Potential Pitfalls of the Mx1-Cre System: Implications for Experimental Modeling of Normal and Malignant Hematopoiesis. *Stem Cell Reports* **7**, 11–18 (2016).
269. Laurenti, E. *et al.* CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell* **16**, 302–313 (2015).
270. Ito, K. *et al.* Non-catalytic Roles of Tet2 Are Essential to Regulate Hematopoietic Stem and Progenitor Cell Homeostasis. *Cell Rep.* **28**, 2480–2490 (2019).
271. Akada, H., Akada, S., Hutchison, R. E. & Mohi, G. Loss of wild-type Jak2

allele enhances myeloid cell expansion and accelerates myelofibrosis in Jak2V617F knock-in mice. *Leukemia* **28**, 1627–1635 (2014).
272. Cimmino, L. *et al.* Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression. *Cell* **170**, 1079–1095 (2017).
273. Miller, C. L. *et al.* Studies of W mutant mice provide evidence for alternate mechanisms capable of activating hematopoietic stem cells. in *Experimental Hematology* (1996).
274. Michel, A. *et al.* Mast Cell–deficient Kit W-sh "Sash" Mutant Mice Display Aberrant Myelopoiesis Leading to the Accumulation of Splenocytes That Act as Myeloid-Derived Suppressor Cells. *J. Immunol.* **190**, 5534–5544 (2013).
275. Ishii, T. *et al.* Behavior of CD34+ cells isolated from patients with polycythemia vera in NOD/SCID mice. *Exp. Hematol.* **35**, 1633–1640 (2007).
276. Delhommeau, F. *et al.* Mutation in *TET2* in Myeloid Cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
277. Dunbar, A., Nazir, A. & Levine, R. Overview of transgenic mouse models of myeloproliferative neoplasms (MPNs). *Curr. Protoc. Pharmacol.* **2017**, (2017).
278. Challen, G. A., Boles, N., Lin, K. K. Y. & Goodell, M. A. Mouse hematopoietic stem cell identification and analysis. *Cytom. Part A* **75**, 14–24 (2009).
279. Deleye, L. *et al.* Performance of four modern whole genome amplification methods for copy number variant detection in single cells. *Sci. Rep.* **7**, (2017).
280. Krumlauf, R. Hox genes in vertebrate development. *Cell* **78**, 191–201 (1994).
281. Dard, A. *et al.* The human HOXA9 protein uses paralog-specific residues of the homeodomain to interact with TALE-class cofactors. *Sci. Rep.* **9**, 1–12 (2019).
282. Wiemels, J. L. *et al.* Site-specific translocation and evidence of postnatal origin of the t(1;19) E2A-PBX1 fusion in childhood acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 15101–15106 (2002).
283. Wong, P., Iwasaki, M., Somervaille, T. C. P., So, C. W. E. & Cleary, M. L. Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. *Genes Dev.* **21**, 2762–2774 (2007).
284. Kroon, E. *et al.* Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.* **17**, 3714–3725 (1998).
285. Dardaei, L., Longobardi, E. & Blasi, F. Prep1 and Meis1 competition for Pbx1 binding regulates protein stability and tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 896–905 (2014).
286. Dardaei, L. *et al.* Tumorigenesis by Meis1 overexpression is accompanied by a change of DNA target-sequence specificity which allows binding to the AP-1 element. *Oncotarget* **6**, 25175–25187 (2015).
287. Mangan, J. K. & Speck, N. A. RUNX1 mutations in clonal myeloid disorders: From conventional cytogenetics to next generation sequencing, A story 40 years in the making. *Crit. Rev. Oncog.* **16**, 77–91 (2011).
288. Ichikawa, M. *et al.* AML1/Runx1 Negatively Regulates Quiescent Hematopoietic Stem Cells in Adult Hematopoiesis. *J. Immunol.* **180**, 4402–4408 (2008).
289. Gerritsen, M. *et al.* RUNX1 mutations enhance self-renewal and block granulocytic differentiation in human in vitro models and primary AMLs. *Blood Adv.* **3**, 320–332 (2019).
290. Willcockson, M. A. *et al.* Runx1 promotes murine erythroid progenitor proliferation and inhibits differentiation by preventing Pu.1 downregulation.

Proc. Natl. Acad. Sci. U. S. A. **116**, 17841–17847 (2019).
291. Sahasrabuddhe, A. A. BMI1: A Biomarker of Hematologic Malignancies. *Biomark. Cancer* **8**, 65–75 (2016).
292. Bommi, P. V., Dimri, M., Sahasrabuddhe, A. A., Khandekar, J. D. & Dimri, G. P. The polycomb group protein BMI1 is a transcriptional target of HDAC inhibitors. *Cell Cycle* **9**, 2663–2673 (2010).
293. Giustacchini, A. *et al.* Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat. Med.* **23**, 692–702 (2017).
294. Schuettpelz, L. G. & Link, D. C. Regulation of Hematopoietic Stem Cell Activity by Inflammation. *Front. Immunol.* **4**, 204 (2013).
295. King, K. Y. & Goodell, M. A. Inflammatory modulation of HSCs: viewing the HSC as a foundation for the immune response. *Nat. Rev. Immunol.* **11**, 685–692 (2011).
296. Hawkins, E. D. *et al.* T-cell acute leukaemia exhibits dynamic interactions with bone marrow microenvironments. *Nature* **538**, 518–522 (2016).
297. Petti, A. A. *et al.* A general approach for detecting expressed mutations in AML cells using single cell RNA-sequencing. *Nat. Commun.* **10**, 1–16 (2019).
298. Rodriguez-Meira, A. *et al.* Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing. *Mol. Cell* **73**, 1292–1305 (2019).
299. Lee-Six, H. *et al.* Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).